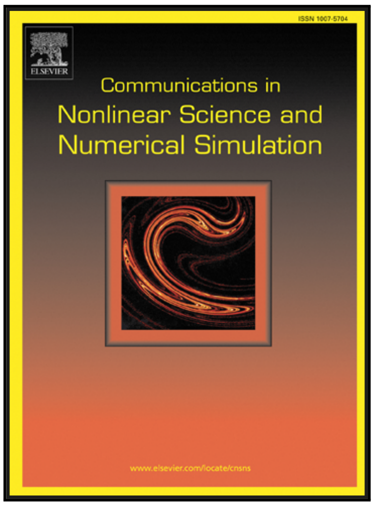
Accepted Manuscript



On the continuation of degenerate periodic orbits via normal form: full dimensional resonant tori

T. Penati, M. Sansottera, V. Danesi

PII: S1007-5704(18)30032-7
DOI: 10.1016/j.cnsns.2018.02.003
Reference: CNSNS 4434

To appear in: *Communications in Nonlinear Science and Numerical Simulation*

Received date: 25 September 2017
Revised date: 2 February 2018
Accepted date: 6 February 2018

Please cite this article as: T. Penati, M. Sansottera, V. Danesi, On the continuation of degenerate periodic orbits via normal form: full dimensional resonant tori, *Communications in Nonlinear Science and Numerical Simulation* (2018), doi: 10.1016/j.cnsns.2018.02.003

Highlights

- the classical problem of the breaking of a completely resonant maximal torus is considered;
- standard averaging methods are not enough when degeneracies occur;
- a new normal form construction is here proposed;
- the normal form algorithm provides high order approximation of degenerate periodic orbits;
- continuation can be then obtained via Newton-Kantorovich scheme.

On the continuation of degenerate periodic orbits via normal form: full dimensional resonant tori

T. Penati*[1], M. Sansottera[1], and V. Danesi[1]

[1]Department of Mathematics, University of Milan, via Saldini 50, 20133 — Milan, Italy.

February 7, 2018

Abstract

We reconsider the classical problem of the continuation of degenerate periodic orbits in Hamiltonian systems. In particular we focus on periodic orbits that arise from the breaking of a completely resonant maximal torus. We here propose a suitable normal form construction that allows to identify and approximate the periodic orbits which survive to the breaking of the resonant torus. Our algorithm allows to treat the continuation of approximate orbits which are at leading order degenerate, hence not covered by classical averaging methods. We discuss possible future extensions and applications to localized periodic orbits in chains of weakly coupled oscillators.

Keywords: normal form construction, completely resonant tori, Hamiltonian perturbation theory, periodic orbits.

1 Introduction

We consider a canonical system of differential equations with Hamiltonian

$$H(I,\varphi,\varepsilon) = H_0(I) + \varepsilon H_1(I,\varphi) + \varepsilon^2 H_2(I,\varphi) + \dots \ , \tag{1}$$

where $I \in \mathcal{U} \subset \mathbb{R}^n$, $\varphi \in \mathbb{T}^n$ are action-angle variables and ε is a small perturbative parameter. The unperturbed system, H_0, is clearly integrable and the orbits, lying on invariant tori, are generically quasi-periodic. Besides, if the unperturbed frequencies satisfy resonance relations, one has periodic orbits on a dense set of resonant tori.

The KAM theorem ensures the persistence of a set of large measure of quasi-periodic orbits, lying on nonresonant tori, for the perturbed system, if ε is small enough and a suitable nondegeneracy condition for H_0 is satisfied.

Instead, considering a resonant torus, when a perturbation is added such a torus is generically destroyed and only a finite number of periodic orbits are expected to survive. The location and stability of the continued periodic orbits are determined by a theorem of Poincaré [35, 36], who approached the problem locally: with an averaging method, he was able to select those isolated unperturbed solutions which, under a suitable nondegeneracy

*Corresponding author, email to tiziano.penati@unimi.it

condition (nowadays called Poincaré nondegeneracy condition), can be continued by means of an implicit function theorem. A modern approach has been developed in the seventies by Weinstein [41] and Moser [28] using bifurcation techniques, turning the problem to the investigation of critical points of a functional on a compact manifold. Actually, the number of critical points can be estimated from below with geometrical methods, like Morse theory. The drawback lies in the fact that the method is not at all constructive, thus it does not permit the localization of the periodic orbits on the torus. In the same spirit, variational methods which make use of the mountain pass theorem were developed some years later by Fadell and Rabinowitz, under different hypotheses (see Chapter 1 in [4] for a simplified exposition of this result). More recently, the problem of continuation of degenerate periodic orbits in nearly integrable Hamiltonian systems using perturbation techniques has been studied in [25, 40]. On the other hand, from the early nineties great attention has been devoted to the generalization of Poincaré's result to partially resonant tori, where the unperturbed torus is foliated by quasi-periodic orbits, since the number of resonances is strictly less than $n-1$. In this case, the starting point still consists in looking for nondegenerate critical points of the perturbation averaged over the unperturbed quasi-periodic solution; however, the presence of more than a single frequency requires the assumption of additional hypotheses, which allow to implement suitable versions of the KAM scheme. Along this line, first results were due to Treshchev [39], Cheng [6], Li and Yi [24]. Recently, these results have been successfully extended to multiscale nearly integrable Hamiltonian systems, where the integrable part of the Hamiltonian $H_0(I, \varepsilon)$, properly involves several time scales, see, e.g., [42,43]. All the quoted works deal with the case where the unperturbed invariant torus is degenerate due to resonances among its frequencies. Instead, we remark that the problems of existence of invariant tori of dimension less than the number of degrees of freedom in weakly perturbed Hamiltonian system, i.e., the extension to lower dimensional tori of the classical KAM theory, has been widely investigated by many authors, see, e.g., [9,23,26,27,29,45–47] in a general abstract framework, and [5,7,8,14,17,38] for more recent problems mainly emerging in Celestial Mechanics.

In this paper we follow the line traced by Poincaré and deal with those cases when the nondegeneracy condition is not fulfilled. In particular, under a twist-like condition of the form (4) (see, e.g., [3]) and analytic estimates of the perturbation (5), we develop an original normal form scheme, inspired by a recent completely constructive proof of the classical Lyapunov theorem on periodic orbits [13], which allows to investigate the continuation of degenerate periodic orbits. Precisely, first we identify possible candidates for the continuation via normal form, then we prove the existence of a unique solution by using the Newton-Kantorovich method.

Remark 1.1 *Let us anticipate a crucial difference with respect to the KAM normal form algorithm: generically, our normal form procedure turns out to be divergent. Actually, a moment's thought suggests that looking for a convergent normal form which is valid for all possible periodic orbits is too much to ask. The idea is that a suitably truncated normal form allows to produce the approximated periodic orbits and the continuation can be performed via contraction or with a further convergent normal form around a selected periodic orbit.*

Remark 1.2 *It is worth mentioning that the idea of performing a finite number of KAM-like steps in order to remove some degeneracy in the continuation procedure is obviously not new, see, e.g., [17, 42, 43] concerning the continuation of quasi-periodic orbits on resonant tori for a class of multiscale nearly integrable Hamiltonian systems. In these works a finite number of* *preliminary KAM steps are performed in order to push the perturbation to a sufficiently high order in ε, before applying a standard convergent scheme.*

The strength of the present perturbative algorithm is at least twofold. First, it provides a way to construct approximate periodic solutions at any desired order in ε, thus going beyond the average approximation mostly used in the literature. One of the few results which represents an improvement with respect to the usual average method is the one claimed in [25], where a criterion for the existence of periodic orbits on completely degenerate resonant tori is proved. In that work the authors, by means of a standard Lindstedt expansion as the original works of Poincaré, are able to push the perturbation scheme at second order in the small parameter ε. However, the possibility to provide a criterion for the continuation, although remarkable, is a consequence of the restriction to completely degenerate cases, like when the Fourier expansion of H_1 with respect to the angle variables does not include a certain resonance class. In this way, all the partial degeneracies are excluded. Such a limitation is overcome by the normal form that we propose: indeed, by being able to deal with any degree of degeneracy, it results more general (also in terms of order of accuracy), thus including also the above mentioned result.

The formal scheme itself has also a second relevant aspect. Since this approximation is given by a recursive explicit algorithm, it can be much useful for numerical applications (see, e.g., [11]) and it is independent of the possibility to conclude the proof with a contraction theorem. Furthermore, our approach provides a constructive normal form that can be applied to a sufficiently general class of models; for example, it includes nonlinear Hamiltonian lattices with next-to-nearest neighbor interactions, such as

$$H = \sum_{j\in\mathcal{J}} \frac{y_j^2}{2} + \sum_{j\in\mathcal{J}} V(x_j) + \varepsilon \sum_{l=1}^{r} \sum_{j\in\mathcal{J}} W(x_{j+l} - x_j) \ ;$$

where $V(x)$ is the potential of an anharmonic oscillator which allows for action variable (at least locally, like the Morse potential), and $W(x)$ represents a generic next-to-nearest neighbour (possibly linear) interaction, with r the maximal range of the interaction. In this class of nearly integrable Hamiltonian lattices, the possibility to generalize the formal scheme to lower dimensional tori would represent a remarkable breakthrough in the investigation of degenerate phase-shift multibreathers and vortexes in one and two dimensional lattices (see, e.g., [1, 2, 21, 22, 30–33]). The extension to lower dimensional tori, that represents the natural continuation of the present work, will be also useful in problems emerging in Celestial Mechanics, where the persistence of nonresonant lower dimensional tori has been proved with similar techniques, see, e.g., [14, 38].

In the present work we focus on resonant maximal tori in order to reduce the technical difficulty to a minimum and concentrate on the novelty of the normal form scheme.

1.1 Outline of the algorithm and statement of the main results

Consider a completely resonant maximal torus of H_0 with unperturbed frequencies

$$\hat{\omega}(I) = \frac{\partial H_0}{\partial I} \ , \quad \text{such that} \quad \hat{\omega}(I) = \omega k \ ,$$

where $\omega \in \mathbb{R}$ and $k \in \mathbb{Z}^n$. This corresponds to a suitable choice of the actions $I = I^*$ with non-vanishing components. From now on, without affecting the generality of the result, we will assume $k_1 = 1$: this will simplify the interpretation of the new variables $\hat{q}, \hat{p}$ that we are going to introduce in a while.

Expanding (1) in power series of the translated actions $J = I - I^*$, one has

$$
\begin{aligned}
H^{(0)} = \langle \hat{\omega}, J \rangle &+ f_4^{(0,0)}(J) + \sum_{l>2} f_{2l}^{(0,0)}(J) \\
&+ f_0^{(0,1)}(\varphi) + f_2^{(0,1)}(J, \varphi) \\
&+ \sum_{s>1} f_0^{(0,s)}(\varphi) + \sum_{s>1} f_2^{(0,s)}(J, \varphi) + \\
&+ \sum_{s>0} \sum_{l>1} f_{2l}^{(0,s)}(J, \varphi) \ ,
\end{aligned}
$$

where $f_{2l}^{(0,s)}$ is a homogeneous polynomial of degree l in J and it is a function of order $\mathcal{O}(\varepsilon^s)$.

Remark 1.3 *The decision to tie the index* $2l$ *to terms of degree* l *in* J *is due to the future extension of the work to lower dimensional tori. Indeed, in that case the transversal directions will be described in cartesian variables, thus the actions will count for two in the total degree. This is also in agreement with the notation adopted in [14].*

Remark 1.4 *The Hamiltonian* (1) *in most applications has only linear terms in the small parameter* ε*, namely* $H_{l \geq 2} \equiv 0$*. Nevertheless, we already consider the general case where the perturbation is analytic in the small parameter. Indeed, as it will be clear from the normal form procedure, starting from the first normalization step we immediately introduce the whole series expansion in* ε*.*

We define the $(n-1)$-dimensional resonant module

$$
\mathcal{M}_\omega = \Big\{ h \in \mathbb{Z}^n : \langle \hat{\omega}, h \rangle = 0 \Big\}
$$

and introduce the resonant variables $\hat{p}$,$\hat{q}$ in place of J, φ. In particular, the pair of conjugate variables $\hat{p}_1, \hat{q}_1$ describes the periodic orbit, while the pairs $\hat{p}_j, \hat{q}_j$, $j = 2, \ldots, n$, represent the transverse directions. The canonical change of coordinates is built with an unimodular matrix (see Lemma 2.10 in [12]) which shows that[1] the new angles $\hat{q}_j$, $j = 2, \ldots, n$, are the phase differences with respect to the *true* angle of the periodic orbit, $\hat{q}_1$, and that $\hat{p}_1$ is given by $\hat{p}_1 = \langle k, J \rangle$.

Introducing the convenient notations $\hat{p} = (p_1, p)$, $\hat{q} = (q_1, q)$ with $p_1 = \hat{p}_1$, $p = (\hat{p}_2, \ldots, \hat{p}_n)$ and correspondingly for q_1 and q, the Hamiltonian can be written in the form

$$
\begin{aligned}
H^{(0)} = \omega p_1 &+ f_4^{(0,0)}(p_1, p) + \sum_{l>2} f_{2l}^{(0,0)}(p_1, p) \\
&+ f_0^{(0,1)}(q_1, q) + f_2^{(0,1)}(p_1, p, q_1, q) \\
&+ \sum_{s>1} f_0^{(0,s)}(q_1, q) + \sum_{s>1} f_2^{(0,s)}(p_1, p, q_1, q) \\
&+ \sum_{s>0} \sum_{l>1} f_{2l}^{(0,s)}(p_1, p, q_1, q)
\end{aligned} \tag{2}
$$

[1]This follows from the assumption $k_1 = 1$. Indeed, in this case that the resonant vector defining the phase differences $\hat{q}_j = k_j \phi_1 - \phi_j$ are a basis for the resonant modulus $\mathcal{M}_\omega$.

where $f_{2l}^{(0,s)}$ is a homogeneous polynomial of degree l in $\hat{p}$ and it is a function of order $\mathcal{O}(\varepsilon^s)$.

Wince we aim to continue a generic unperturbed periodic orbit $p_1 = 0$, $q_1 = q_1(0) + \omega t$, $p = 0$, $q = q^*$, we look for a normal form which is able to select those phase shifts, q^*, which represent good candidates for continuation. The Hamiltonian is said to be in normal form up to order r if the constant and linear terms in the actions are averaged (up to order r) with respect to the fast angle, q_1, and if, for a fixed but arbitrary q^*, the linear terms in the action, evaluated at $q = q^*$, vanishes identically.

In order to give a precise statement we need to introduce the mathematical framework. We consider the extended complex domains $\mathcal{D}_{\rho,\sigma} = \mathcal{G}_\rho \times \mathbb{T}^n_\sigma$, defined as

$$\mathcal{G}_\rho = \left\{ \hat{p} \in \mathbb{C}^n : \max_{1\le j\le n} |\hat{p}_j| < \rho \right\} ,$$
$$\mathbb{T}^n_\sigma = \left\{ \hat{q} \in \mathbb{C}^n : \operatorname{Re} \hat{q}_j \in \mathbb{T},\ \max_{1\le j\le n} |\operatorname{Im} \hat{q}_j| < \sigma \right\} ,$$

and introduce the distinguished classes of functions $\mathcal{P}_{2l}$, with integers l, which can be written as a Fourier-Taylor expansion

$$g(\hat{p}, \hat{q}) = \sum_{\substack{i\in\mathbb{N}^n\\ |i|=l}} \sum_{k\in\mathbb{Z}^n} g_{i,k}\, \hat{p}^i e^{\mathbf{i}\langle k, \hat{q}\rangle} , \tag{3}$$

with coefficients $g_{i,k} \in \mathbb{C}$. We also set $\mathcal{P}_{-2} = \{0\}$.

For a generic analytic function $g \in \mathcal{P}_{2l}$, $g : \mathcal{D}_{\rho,\sigma} \to \mathbb{C}$, we define the weighted Fourier norm

$$\|g\|_{\rho,\sigma} = \sum_{\substack{i\in\mathbb{N}^n\\ |i|=l}} \sum_{k\in\mathbb{Z}^n} |g_{i,k}| \rho^l e^{|k|\sigma} .$$

Hereafter, we use the shorthand notation $\|\cdot\|_\alpha$ for $\|\cdot\|_{\alpha(\rho,\sigma)}$.

We state here our main result concerning the normal form.

Proposition 1.1 *Consider a Hamiltonian $H^{(0)}$ expanded as in (2) that is analytic in a domain $\mathcal{D}_{\rho,\sigma}$. Let us assume that*

(a) there exists a positive constant m such that for every $v \in \mathbb{R}^n$ one has

$$m \sum_{i=1}^n |v_i| \le \sum_{i=1}^n \Big| \sum_{j=1}^n C_{ij} v_j \Big| , \quad \textit{where} \quad C_{ij} = \frac{\partial^2 f_4^{(0,0)}}{\partial \hat{p}_i \partial \hat{p}_j} ; \tag{4}$$

(b) the terms appearing in the expansion of the Hamiltonian satisfy

$$\|f_{2l}^{(0,s)}\|_1 \le \frac{E}{2^{2l}} \varepsilon^s , \quad \textit{with} \quad E > 0. \tag{5}$$

Then, for every positive integer r there is a positive ε_r^* such that for $0 \le \varepsilon < \varepsilon_r^*$ there exists an analytic canonical transformation $\Phi^{(r)}$ satisfying

$$\mathcal{D}_{\frac{1}{4}(\rho,\sigma)} \subset \Phi^{(r)}\left(\mathcal{D}_{\frac{1}{2}(\rho,\sigma)}\right) \subset \mathcal{D}_{\frac{3}{4}(\rho,\sigma)} \tag{6}$$

such that the Hamiltonian $H^{(r)} = H^{(0)} \circ \Phi^{(r)}$ is in normal form up to order r, namely

$$
\begin{aligned}
H^{(r)}(p_1, p, q_1, q; q^*) = \omega p_1 + f_4^{(r,0)}(p_1, p) + \sum_{l>2} f_{2l}^{(r,0)}(p_1, p) \\
+ \sum_{s=1}^{r} f_0^{(r,s)}(q; q^*) + \sum_{s=1}^{r} f_2^{(r,s)}(p_1, p, q; q^*) \\
+ \sum_{s>r} f_0^{(r,s)}(q_1, q; q^*) + \sum_{s>r} f_2^{(r,s)}(p_1, p, q_1, q; q^*) \\
+ \sum_{s>0} \sum_{l>1} f_{2l}^{(r,s)}(p_1, p, q_1, q; q^*) \ ,
\end{aligned}
\tag{7}
$$

where q^* is a fixed but arbitrary parameter and $f_{2l}^{(r,s)} \in \mathcal{P}_{2l}$ is a function of order $\mathcal{O}(\varepsilon^s)$. Moreover, for $q = q^*$ one has

$$
\sum_{s=1}^{r} f_2^{(r,s)}(p_1, p, q^*; q^*) = 0 \ . \tag{8}
$$

The Hamilton equations associated to the truncated normal form, i.e., neglecting term of order $\mathcal{O}(\varepsilon^{r+1})$, once evaluated at $(\hat{p} = 0, q = q^*)$, read

$$
\dot{p}_1 = 0 \ , \qquad \dot{q}_1 = \omega \ , \qquad \dot{p} = -\sum_{s=1}^{r} \nabla_q f_0^{(r,s)} \ , \qquad \dot{q} = 0 \ .
$$

Hence, if

$$
\sum_{s=1}^{r} \nabla_q f_0^{(r,s)} \big|_{q=q^*} = 0 \ , \tag{9}
$$

then $p_1 = 0$, $q_1 = q_1(0)$, $p = 0$, $q = q^*$ is the initial datum of a periodic orbit with frequency ω for the truncated normal form. Considering the whole system given by $H^{(r)}$, the initial datum provides an *approximate* periodic orbit with frequency ω, which turns out to be a relative equilibrium of the truncated Hamiltonian. In order to provide a precise definition of *approximate periodic orbit* we introduce the T-period map $\Upsilon : \mathcal{U}(q^*, 0) \subset \mathbb{R}^{2n-1} \to \mathcal{V}(q^*, 0) \subset \mathbb{R}^{2n-1}$, a smooth function of the $2n-1$ variables $(q, \hat{p})$, parametrized by the initial phase $q_1(0)$ and the small parameter ε, precisely

$$
\Upsilon(q(0), \hat{p}(0); \varepsilon, q_1(0)) = \begin{pmatrix} F(q(0), \hat{p}(0); \varepsilon, q_1(0)) \\ G(q(0), \hat{p}(0); \varepsilon, q_1(0)) \end{pmatrix} = \begin{pmatrix} \hat{q}(T) - \hat{q}(0) - \Lambda T \\ \frac{1}{\varepsilon}(p(T) - p(0)) \end{pmatrix} \ , \tag{10}
$$

with $\Lambda = (\omega, 0) \in \mathbb{R}^n$. The map Υ represents the T-flow of the $n-1$ actions p and of the n angles $\hat{q}$ for the Hamiltonian $H^{(r)}$.

Let us stress that $p_1 = 0$, $q_1 = q_1$, $p = 0$, $q = q^*$ corresponds to a periodic orbit for the truncated normal form, thus it is evident that $\Upsilon(q^*, 0; \varepsilon, q_1(0))$ is of order[2] $\mathcal{O}(\varepsilon^r)$. Thus, a true periodic orbit, close to the approximate one, is identified by an initial datum $(q^*_{\text{p.o.}}, \hat{p}_{\text{p.o.}}) \in \mathcal{U}(0, q^*)$ such that

$$
\Upsilon(q^*_{\text{p.o.}}, \hat{p}_{\text{p.o.}}; \varepsilon, q_1(0)) = 0 \ .
$$

[2]The actions p have been rescaled by ε in Υ, hence only G is of order $\mathcal{O}(\varepsilon^{r+1})$ while F is of order $\mathcal{O}(\varepsilon^r)$.

In order to prove the existence of a unique solution $q^* = q^*_{\rm p.o.}$, $\hat{p} = \hat{p}_{\rm p.o.}$, $q_1 = q_1(0)$, close enough to the approximate one, we apply the Newton-Kantorovich algorithm. Therefore we need to ensure that the Jacobian matrix (with respect to the initial datum)

$$M(\varepsilon) = D_{\hat{p}(0),q(0)}\Upsilon(q^*,0;\varepsilon,q_1(0)) \tag{11}$$

is invertible and its eigenvalues are not too small with respect to ε^r.

We state here the main result concerning the continuation of the periodic orbits

Theorem 1.1 *Consider the map* Υ *defined in* (10) *in a neighbourhood of the torus* $\hat{p} = 0$ *and let* $(q^*(\varepsilon), 0)$*, with* $q^*(\varepsilon)$ *satisfying* (9)*, an approximate zero of* Υ*, namely*

$$\|\Upsilon(q^*(\varepsilon),0;\varepsilon,q_1(0))\| \le C_1\varepsilon^r \ ,$$

where C_1 *is a positive constant just depending on* $\mathcal{U}$*. Assume that the matrix* $M(\varepsilon)$ *defined in* (11) *is invertible and its eigenvalues satisfy*

$$|\lambda| \ge \varepsilon^\alpha \ , \qquad \textit{for} \quad \lambda \in {\rm spec}(M(\varepsilon)) \quad \textit{with} \quad 2\alpha < r \ . \tag{12}$$

Then, there exist $C_0 > 0$ *and* $\varepsilon^* > 0$ *such that for any* $0 \le \varepsilon < \varepsilon^*$ *there exists a unique* $(q^*_{\rm p.o.}(\varepsilon), \hat{p}_{\rm p.o.}(\varepsilon)) \in \mathcal{U}$ *which solves*

$$\Upsilon(q^*_{\rm p.o.},\hat{p}_{\rm p.o.};\varepsilon,q_1(0)) = 0 \ , \qquad \left\|(q^*_{\rm p.o.},\hat{p}_{\rm p.o.}) - (q^*,0)\right\| \le C_0\varepsilon^{r-\alpha} \ . \tag{13}$$

Before entering the technical part of the paper, let us add some more considerations. First, as already remarked, the above Theorem generalizes an old and classical result by Poincaré, whose idea was to average the perturbation H_1 with respect to the flow of the unperturbed periodic solution, where only the fast angle q_1 rotates. The candidates q^* for continuation were the nondegenerate relative extrema on the torus $\mathbb{T}^{n-1}$ of the averaged Hamiltonian $\langle H_1\rangle_{q_1}$, namely

$$\nabla_q\langle H_1\rangle_{q_1} = 0 \ , \qquad |D^2_q\langle H_1\rangle_{q_1}| \ne 0 \ .$$

The result of Poincaré actually corresponds to the construction of the first order normal form together with a nondegeneracy assumption on the ε-independent version of (9), precisely

$$\nabla_q f_0^{(1,1)} = 0 \ , \qquad |D^2_q f_0^{(1,1)}| \ne 0 \ . \tag{14}$$

In such a case, due to the simplified form of Υ, the solution $(\hat{p}_{\rm p.o.}, q^*_{\rm p.o.})$ can be obtained via implicit function theorem in a neighborhood of the approximate initial datum $(0, q^*)$, q^* being a solution of the first of (14), independent of ε. Hence, our high-order normal form construction becomes a necessary way in order to deal with *degenerate cases*, where solutions of (14) are not isolated and appear as d-parameter families, thus leading to $|D^2_q f_0^{(1,1)}| = 0$.

For instance, in the application presented in Section 4, the solutions of (14) show up as one parameter families $q^*(s)$. Actually, solving (9) (with $r \ge 2$) in place of (14) allows to isolate true candidates for the continuation. Let us also remark that our scheme provides a refined averaged Hamiltonian which allows to treat the totally degenerate case, i.e., $\nabla_q f_0^{(1,1)} \equiv 0$. In particular, the results presented in [25] by means of Lindstedt perturbation scheme can be obtained as special cases.

The paper is organized as follows. In Section 2 we detail the normal form algorithm together with the quantitative estimates. The proof of Theorem 1.1 is reported in Section 3.

Section 4 provides a simplified version of Theorem 1.1, namely Theorem 4.1, for one parameter families of solutions of (14), under the assumption that only the second normal form step is enough to improve the accuracy of the approximate periodic orbit. Moreover, a pedagogical example inspired by the problem of degenerate vortexes in a squared lattice dNLS model is presented at the end of Section 4. Appendices A and B include the technicalities related to the normal form estimates and the Newton-Kantorovich method, respectively.

Let us remark that, since one expects that two normalization steps allow to deal with one-parameter families of potential periodic orbits (as in the example reported in Section 4), we explicitly report in Appendix A all the quantitative estimates for the first two normalization steps. This will allow to directly exploit these estimates in future applications.

2 Normal formal algorithm and analytical estimates

This Section is devoted to the formal algorithm that takes a Hamiltonian (2) and brings it into normal form up to an arbitrary, but finite, order r. We include all the (often tedious) formulæ that will be used in order to estimate the terms appearing in the normalization process. We use the formalism of Lie series and Lie transforms (see, e.g., [16] and [12] for a self-consistent introduction).

The transformation at step r is generated via composition of two Lie series of the form

$$\exp(L_{\chi_2^{(r)}}) \circ \exp(L_{\chi_0^{(r)}}) \ ,$$

where

$$\chi_0^{(r)} = X_0^{(r)} + \langle \zeta^{(r)}, \varphi \rangle \ , \tag{15}$$

with $\zeta^{(r)} \in \mathbb{R}^n$ and $X_0^{(r)} \in \mathcal{P}_0$, $\chi_2^{(r)} \in \mathcal{P}_2$ are of order $\mathcal{O}(\varepsilon^r)$. Here, as usual, we denote by $L_g\cdot$ the Poisson bracket $\{\cdot, g\}$. The functions $\chi_0^{(r)}$ and $\chi_2^{(r)}$ are unknowns to be determined so that the transformed Hamiltonian is in normal form up to order r.

The relevant algebraic property of the $\mathcal{P}_\ell$ classes of function is stated by the following

Lemma 2.1 *Let $f \in \mathcal{P}_{s_1}$ and $g \in \mathcal{P}_{s_2}$, then $\{f, g\} \in \mathcal{P}_{s_1+s_2-2}$.*

The straightforward proof is left to the reader.

The starting Hamiltonian has the form

$$\begin{aligned} H^{(0)} = \omega p_1 + \sum_{s\geq 0} \sum_{l>1} f_{2l}^{(0,s)} \\ + \sum_{s\geq 1} f_0^{(0,s)} + \sum_{s\geq 1} f_2^{(0,s)} \ , \end{aligned} \tag{16}$$

where $f_{2l}^{(0,s)} \in \mathcal{P}_{2l}$ and is of order $\mathcal{O}(\varepsilon^s)$.

We now describe the generic r-th normalization step, starting from the Hamiltonian in normal form up to order $r-1$, $H^{(r-1)}$, namely

$$\begin{aligned} H^{(r-1)} = \omega p_1 &+ \sum_{s<r} f_0^{(r-1,s)} + \sum_{s<r} f_2^{(r-1,s)} \\ &+ f_0^{(r-1,r)} + f_2^{(r-1,r)} \\ &+ \sum_{s>r} f_0^{(r-1,s)} + \sum_{s>r} f_2^{(r-1,s)} \\ &+ \sum_{s\geq 0} \sum_{l>1} f_{2l}^{(r-1,s)} \ , \end{aligned} \tag{17}$$

where $f_{2l}^{(r-1,s)} \in \mathcal{P}_{2l}$ is of order $\mathcal{O}(\varepsilon^s)$; $f_0^{(r-1,s)}$ and $f_2^{(r-1,s)}$ for $1 \leq s < r$ are in normal form.

2.1 First stage of the normalization step

Our aim is to put the term $f_0^{(r-1,r)}$ in normal form and to keep fixed the harmonic frequencies of the selected resonant torus. We determine the generating function $\chi_0^{(r)} = X_0^{(r)} + \langle \zeta^{(r)}, \hat{q} \rangle$ by solving the homological equations

$$L_{X_0^{(r)}} \omega p_1 + f_0^{(r-1,r)} = \langle f_0^{(r-1,r)} \rangle_{q_1} \ ,$$
$$L_{\langle \zeta^{(r)}, \hat{q} \rangle} f_4^{(0,0)} + \left\langle f_2^{(r-1,r)} \Big|_{q=q^*} \right\rangle_{q_1} = 0 \ .$$

Considering the Taylor-Fourier expansion

$$f_0^{(r-1,r)}(\hat{q}) = \sum_k c_{0,k}^{(r-1,r)} \exp(\mathbf{i}\langle k, \hat{q} \rangle) \ ,$$

we readily get

$$X_0^{(r)}(\hat{q}) = \sum_{k_1 \neq 0} \frac{c_{0,k}^{(r-1,r)}}{\mathbf{i} k_1 \omega} \exp(\mathbf{i}\langle k, \hat{q} \rangle) \ .$$

The translation vector, $\zeta^{(r)}$, is determined by solving the linear system

$$\sum_j C_{ij} \zeta_j^{(r)} = \frac{\partial}{\partial \hat{p}_i} \left\langle f_2^{(r-1,r)} \Big|_{q=q^*} \right\rangle_{q_1} \ . \tag{18}$$

This translation, which involves the linear term in the actions $f_2^{(r-1,r)}$, allows to keep fixed the frequency ω and kills the small transversal frequencies in the angles q.

The transformed Hamiltonian is computed as

$$H^{(\mathrm{I};r-1)} = \exp\left(L_{\chi_0^{(r)}} \right) H^{(r-1)}$$

and has a form similar to (17), precisely

$$\begin{aligned} H^{(\mathrm{I};r-1)} &= \exp\Big(L_{\chi_0^{(r)}}\Big) H^{(r-1)} = \\ &= \omega p_1 + \sum_{s<r} f_0^{(\mathrm{I};r-1,s)} + \sum_{s<r} f_2^{(\mathrm{I};r-1,s)} \\ &\quad + f_0^{(\mathrm{I};r-1,r)} + f_2^{(\mathrm{I};r-1,r)} \\ &\quad + \sum_{s>r} f_0^{(\mathrm{I};r-1,s)} + \sum_{s>r} f_2^{(\mathrm{I};r-1,s)} \\ &\quad + \sum_{s\geq 0}\sum_{l>1} f_{2l}^{(\mathrm{I};r-1,s)} \ . \end{aligned} \tag{19}$$

The functions $f_{2l}^{(\mathrm{I};r-1,s)}$ are recursively defined as

$$\begin{aligned} f_0^{(\mathrm{I};r-1,r)} &= \big\langle f_0^{(r-1,r)} \big\rangle_{q_1} \ , \\ f_2^{(\mathrm{I};r-1,r)} &= f_2^{(r-1,r)} - \big\langle f_2^{(r-1,r)}(q^*) \big\rangle_{q_1} + L_{X_0^{(r)}} f_4^{(0,0)} \ , \\ f_{2l}^{(\mathrm{I};r-1,s)} &= \sum_{j=0}^{\lfloor s/r \rfloor} \frac{1}{j!} L^j_{\chi_0^{(r)}} f_{2l+2j}^{(r-1,s-jr)} \ , \qquad \text{for } l=0,1,\ s\neq r \ , \\ & \qquad\qquad\qquad\qquad\qquad\qquad \text{or } l\geq 2,\ s\geq 0 \ , \end{aligned} \tag{20}$$

with $f_{2l}^{(\mathrm{I};r-1,s)} \in \mathcal{P}_{2l}$.

2.2 Second stage of the normalization step

We now put $f_2^{(\mathrm{I};r-1,r)}$ in normal form, by averaging with respect to the fast angle q_1. This is necessary in order to avoid small oscillations of q around q^*. We determine the generating function $\chi_2^{(r)}$ by solving the homological equation

$$L_{\chi_2^{(r)}} \omega p_1 + f_2^{(\mathrm{I};r-1,r)} = \Big\langle f_2^{(\mathrm{I};r-1,r)} \Big\rangle_{q_1} \ .$$

Considering again the Taylor-Fourier expansion

$$f_2^{(\mathrm{I};r-1,r)}(\hat{p},\hat{q}) = \sum_{\substack{|l|=1 \\ k}} c_{l,k}^{(\mathrm{I};r-1,r)} \hat{p}^l \exp(\mathbf{i}\langle k,\hat{q}\rangle)$$

we get

$$\chi_2^{(r)}(\hat{p},\hat{q}) = \sum_{\substack{|l|=1 \\ k_1\neq 0}} \frac{c_{l,k}^{(\mathrm{I};r-1,r)} \hat{p}^l \exp(\mathbf{i}\langle k,\hat{q}\rangle)}{\mathbf{i} k_1 \omega} \ .$$

The transformed Hamiltonian is computed as

$$H^{(r)} = \exp\left(L_{\chi_2^{(r)}}\right) H^{(\mathrm{I};r-1)}$$

and is given the form (17), replacing the upper index $r-1$ by r, with

$$
\begin{aligned}
f_2^{(r,r)} &= \langle f_2^{(\mathrm{I};r-1,r)}\rangle_{q_1} \ , \\
f_2^{(r,jr)} &= \frac{1}{(j-1)!} L_{\chi_2^{(r)}}^{j-1} \left(\frac{1}{j} \langle f_2^{(\mathrm{I};r-1,r)}\rangle_{q_1} + \frac{j-1}{j} f_2^{(\mathrm{I};r-1,r)} \right) \\
&\quad + \sum_{j=0}^{\lfloor s/r\rfloor-2} \frac{1}{j!} L_{\chi_2^{(r)}}^{j} f_2^{(\mathrm{I};r-1,s-jr)} \ , \\
f_{2l}^{(r,s)} &= \sum_{j=0}^{\lfloor s/r\rfloor} \frac{1}{j!} L_{\chi_2^{(r)}}^{j} f_{2l}^{(\mathrm{I};r-1,s-jr)} \qquad \begin{aligned} &\text{for } l=0,\ s\geq 0\ , \\ &\text{or } l=1,\ s\neq jr\ , \\ &\text{or } l\geq 2,\ s\geq 0\ . \end{aligned}
\end{aligned}
\tag{21}
$$

2.3 Analytic estimates

In order to translate our formal algorithm into a recursive scheme of estimates on the norms of the various functions, we need to introduce a sequence of restrictions of the domain so as to apply Cauchy's estimate. Having fixed $d\in\mathbb{R}$, $0<d\leq 1/4$, we consider a sequence $\delta_{r\geq 1}$ of positive real numbers satisfying

$$
\delta_{r+1}\leq\delta_r \ , \quad \sum_{r\geq 1}\delta_r\leq\frac{d}{2} \ ; \tag{22}
$$

thus the sequence δ_r has to satisfy the inequality $\delta_r<C/r$ for some $r>\overline{r}$ and $C\in\mathbb{R}$. Moreover, we introduce a further sequence $d_{r\geq 0}$ of real numbers recursively defined as

$$
d_0=0\ , \quad d_r=d_{r-1}+2\delta_r \ . \tag{23}
$$

In order to precisely state the iterative Lemma, we need to introduce the quantities Ξ_r, parametrized by the index r, as

$$
\Xi_r=\max\left(1,\frac{E}{\omega\delta_r^2\rho\sigma}+\frac{eE}{4m\delta_r\rho^2},2+\frac{E}{2e\omega\delta_r\rho\sigma},\frac{E}{4\omega\delta_r^2\rho\sigma}\right) \ . \tag{24}
$$

Following the approach described in [10], the number of terms generated recursively by formulæ (20) and (21) is controlled by the two sequences $\{\nu_{r,s}\}_{r\geq 0\,,\,s\geq 0}$ and $\{\nu_{r,s}^{(\mathrm{I})}\}_{r\geq 1\,,\,s\geq 0}$ of integer numbers that are recursively defined as

$$
\begin{aligned}
\nu_{0,s} &= 1 && \text{for } s\geq 0\,, \\
\nu_{r,s}^{(\mathrm{I})} &= \sum_{j=0}^{\lfloor s/r\rfloor} \nu_{r-1,r}^{j}\nu_{r-1,s-jr} && \text{for } r\geq 1\,,\ s\geq 0\,, \\
\nu_{r,s} &= \sum_{j=0}^{\lfloor s/r\rfloor} (3\nu_{r-1,r})^{j}\nu_{r,s-jr}^{(\mathrm{I})} && \text{for } r\geq 1\,,\ s\geq 0\,.
\end{aligned}
\tag{25}
$$

Let us stress that when $s<r$, the above simplify as

$$
\nu_{r,s}^{(\mathrm{I})}=\nu_{r-1,s}\ , \qquad \nu_{r,s}=\nu_{r,s}^{(\mathrm{I})}\ ,
$$

namely

$$\nu_{r,s} = \nu_{r-1,s} = \ldots = \nu_{s,s} \ .$$

Let us introduce the quantities $b(\mathrm{I};r,s,2l)$ and $b(r,s,2l)$ (r being a positive integer, while s and l are non-negative ones) that will be useful to control the exponents of the Ξ_r in the normalization procedure,

$$b(\mathrm{I};r,s,2l) = \begin{cases} s & \text{if } r=1\ ,\\ 0 & \text{if } r\geq 2, s=0\ ,\\ 3s-\left\lfloor\frac{s+r-1}{r}\right\rfloor-\left\lfloor\frac{s+r-2}{r}\right\rfloor-2 & \text{if } r\geq 2, 0<s\leq r, l=0\\ 3s-\left\lfloor\frac{s+r-1}{r}\right\rfloor-\left\lfloor\frac{s+r-2}{r}\right\rfloor-1 & \text{if } r\geq 2, r<s\leq 2r, l=0\\ 3s-\left\lfloor\frac{s+r-1}{r}\right\rfloor-\left\lfloor\frac{s+r-2}{r}\right\rfloor-1 & \text{if } r\geq 2, 0<s\leq r, l=1\\ 3s-\left\lfloor\frac{s+r-1}{r}\right\rfloor-\left\lfloor\frac{s+r-2}{r}\right\rfloor & \text{in the other cases} \end{cases}$$

and

$$b(r,s,2l) = \begin{cases} 0 & \text{if } r>0, s=0\\ 3s-\left\lfloor\frac{s+r-1}{r}\right\rfloor-w_{2l} & \text{if } r=1, s>0\ ,\\ 3s-\left\lfloor\frac{s+r-1}{r}\right\rfloor-\left\lfloor\frac{s+r-2}{r}\right\rfloor-2 & \text{if } r\geq 2, 0<s\leq r, l=0\\ 3s-\left\lfloor\frac{s+r-1}{r}\right\rfloor-\left\lfloor\frac{s+r-2}{r}\right\rfloor-1 & \text{if } r\geq 2, r<s\leq 2r, l=0\\ 3s-\left\lfloor\frac{s+r-1}{r}\right\rfloor-\left\lfloor\frac{s+r-2}{r}\right\rfloor-1 & \text{if } r\geq 2, 0<s\leq r, l=1\\ 3s-\left\lfloor\frac{s+r-1}{r}\right\rfloor-\left\lfloor\frac{s+r-2}{r}\right\rfloor & \text{in the other cases} \end{cases}$$

with $w_0 = 2$, $w_2 = 1$ and $w_{2l} = 0$ for $l \geq 2$.

We are now ready to state the main Lemma collecting the estimates for the generic r-th normalization step of the normal form algorithm.

Lemma 2.2 *Consider a Hamiltonian $H^{(r-1)}$ expanded as in (17). Let $\chi_0^{(r)} = X_0^{(r)} + \langle\zeta^{(r)},\varphi\rangle$ and $\chi_2^{(r)}$ be the generating functions used to put the Hamiltonian in normal form at order r, then one has*

$$\begin{aligned} \|X_0^{(r)}\|_{1-d_{r-1}} &\leq \frac{1}{\omega}\nu_{r-1,r}\Xi_r^{3r-4}E\varepsilon^r\ ,\\ |\zeta^{(r)}| &\leq \frac{1}{4m\rho}\nu_{r-1,r}\Xi_r^{3r-3}E\varepsilon^r\ ,\\ \|\chi_2^{(r)}\|_{1-d_{r-1}-\delta_r} &\leq \frac{1}{\omega}3\nu_{r-1,r}\Xi_r^{3r-3}\frac{E}{4}\varepsilon^r\ . \end{aligned} \tag{26}$$

The terms appearing in the expansion of $H^{(\mathrm{I};r-1)}$ in (19) are bounded as

$$\|f_{2l}^{(\mathrm{I;r-1,s})}\|_{1-d_{r-1}-\delta_r} \leq \nu_{r,s}^{(\mathrm{I})}\Xi_r^{b(\mathrm{I};r,s,2l)}\frac{E}{2^{2l}}\varepsilon^s\ .$$

The terms appearing in the expansion of $H^{(r)}$ in (21) are bounded as

$$\|f_{2l}^{(r,s)}\|_{1-d_r} \leq \nu_{r,s}\Xi_r^{b(r,s,2l)}\frac{E}{2^{2l}}\varepsilon^s\ .$$

The proof of Lemma 2.2 is deferred to Section A.4.1. Besides, some comments about the statement of this Lemma are in order.

Remark 2.1 *The well-known problem of the accumulation of small divisors represents the source of divergence in perturbation processes. In the present work we are considering a completely resonant normal form, thus, if $\omega \neq 0$, the divisors $k_1\omega$ introduced in the solution of the homological equations (see Sections 2.1 and 2.1) cannot become arbitrarily small. In particular, we do not need any strong nonresonance condition on the frequencies. However, the restrictions of the domains due to the Cauchy estimates for derivatives, introduce the small denominators δ_r that actually accumulate to zero and are the responsible for the divergence of the normal form.*

Remark 2.2 *A remarkable technical difference with respect to the analytical estimates of the Kolmogorov theorem is the factor 3 (instead of 2) in the exponents of the Ξ_r^r. This is due to the different terms appearing in our resonant normal form. However, as we do not have any nonresonance condition on the frequency vector ω, the problem of the optimality of the factor appearing in the exponents is not crucial as in other related works, see, e.g., [15] concerning the Schröder-Siegel problem. Given this, our impression is that the factor 3 can be hardly improved.*

2.4 Proof of Proposition 1.1

We give here a sketch of the proof of Proposition 1.1. The proof is based on standard arguments in Lie series theory, that we recall here, referring to, e.g., [10, 14, 37], for more details.

We give an estimate for the canonical transformation. We denote by $(\hat{p}^{(0)}, \hat{q}^{(0)})$ the original coordinates, and by $(\hat{p}^{(r)}, \hat{q}^{(r)})$ the coordinates at step r. We also denote by $\phi^{(r)}$ the canonical transformation mapping $(\hat{p}^{(r)}, \hat{q}^{(r)})$ to $(\hat{p}^{(r-1)}, \hat{q}^{(r-1)})$, precisely

$$\begin{aligned}
\hat{p}^{(r-1)} &= \exp(L_{\chi_0^{(r)}})\hat{p}^{(\mathrm{I},r-1)} = \hat{p}^{(\mathrm{I},r-1)} + \frac{\partial \chi_0^{(r)}}{\partial \hat{q}^{(r-1)}} \ , \\
\hat{p}^{(\mathrm{I},r-1)} &= \exp(L_{\chi_2^{(r)}})\hat{p}^{(r)} = \hat{p}^{(r)} + \sum_{s\geq 1} \frac{1}{s!} L^{s-1}_{\chi_2^{(r)}} \frac{\partial \chi_2^{(r)}}{\partial \hat{q}^{(r)}} \ , \\
\hat{q}^{(r-1)} &= \exp(L_{\chi_2^{(r)}})\hat{q}^{(r)} = \hat{q}^{(r)} - \sum_{s\geq 1} \frac{1}{s!} L^{s-1}_{\chi_2^{(r)}} \frac{\partial \chi_2^{(r)}}{\partial \hat{p}^{(r)}} \ .
\end{aligned}$$

Consider now a sequence of domains $\mathcal{D}_{(3d-d_r)(\rho,\sigma)}$, using Lemma 2.2 we get

$$\begin{aligned}
\left|\hat{p}^{(r-1)} - \hat{p}^{(\mathrm{I},r-1)}\right| &< \left(\frac{1}{\omega e \delta_r \sigma} + \frac{1}{4m\rho}\right) \Xi_r^{3r} \frac{100^r}{20} E\varepsilon^r \ , \\
\left|\hat{p}^{(\mathrm{I},r-1)} - \hat{p}^{(r)}\right| &< \frac{1}{4\omega e \delta_r \sigma} \Xi_r^{3r} \frac{100^r}{20} E\varepsilon^r \sum_{s\geq 1} \left(\frac{1}{\omega \delta_r^2 \rho\sigma} \Xi_r^{3r} \frac{100^r}{20} E\varepsilon^r\right)^{s-1} \ , \\
\left|\hat{q}^{(r-1)} - \hat{q}^{(r)}\right| &< \frac{1}{4\omega \delta_r \rho} \Xi_r^{3r} \frac{100^r}{20} E\varepsilon^r \sum_{s\geq 1} \left(\frac{1}{\omega \delta_r^2 \rho\sigma} \Xi_r^{3r} \frac{100^r}{20} E\varepsilon^r\right)^{s-1} \ .
\end{aligned} \tag{27}$$

Thus if ε is small enough (for a very rough estimate take $\varepsilon < \frac{1}{100\Xi_r^4}$) the series (27) defining the canonical transformation are absolutely convergent in the domain $\mathcal{D}_{(3d-d_{r-1}-\delta_r)(\rho,\sigma)}$, hence analytic. Furthermore, one has the estimates

$$|\hat{p}^{(r-1)} - \hat{p}^{(r)}| < \delta_r \rho \ , \qquad |\hat{q}^{(r-1)} - \hat{q}^{(r)}| < \delta_r \sigma \ .$$

A similar argument applies to the inverse of $\phi^{(r)}$, which is defined as a composition of Lie series generated by $\chi_2^{(r)}$ and $-\chi_0^{(r)}$, thus we get

$$\mathcal{D}_{(3d-d_r)(\rho,\sigma)} \subset \phi^{(r)}\big(\mathcal{D}_{(3d-d_{r-1}-\delta_r)(\rho,\sigma)}\big) \subset \mathcal{D}_{(3d-d_{r-1})(\rho,\sigma)} \ .$$

Consider now the sequence of transformations $\Phi^{(\bar r)} = \phi^{(1)} \circ \ldots \circ \phi^{(\bar r)}$. For $(\hat p^{(r-1)}, \hat q^{(r-1)}) \in \mathcal{D}_{(3d-d_{r-1})(\rho,\sigma)}$ the transformation is clearly analytic and one has

$$|\hat p^{(0)} - \hat p^{(\bar r)}| < \rho \sum_{j=1}^{\bar r} \delta_j \ , \qquad |\hat q^{(0)} - \hat q^{(\bar r)}| < \sigma \sum_{j=1}^{\bar r} \delta_j \ .$$

Setting $d = \frac{1}{4}$ and using (22), one has $\sum_{j\geq 1} \delta_j \leq \frac{d}{2} = \frac{1}{8}$, thus (6) immediately follows. Finally, the estimates for the Hamiltonian in normal form had been already gathered in Lemma 2.2. This concludes the proof of Proposition 1.1.

Remark 2.3 *Since the non convergence of the normalization algorithm represents one of the main points, let us stress that in view of the the definition of* Ξ_r *in* (24) *and of* $\delta_r < C/r$, *one immediately get* $\Xi_r > Cr$, C *being a suitable positive constant. Thus* $\sum_{r>0} \Xi_r^{3r} \varepsilon^r$ *cannot converge for any positive* ε.

3 Proof of Theorem 1.1

In this Section we develop in a more detailed way the strategy used to get Theorem 1.1 from the normal form constructed. We have shown in the previous Section that, by means of a canonical and near the identity change of coordinates, it is possible to give the original Hamiltonian the form (7). We have already stressed in the Introduction the main feature of our construction: if one considers the approximate equations of motion corresponding to the normal form truncated at order $\mathcal{O}(\varepsilon^r)$, when evaluated on $(q = q^*, \hat p = 0)$, they provide a periodic orbit of frequency ω once q^* fulfills the already mentioned equation (9). Generically, for $r \geq 2$, the value q^* would depend continuously on ε, precisely $q^*(\varepsilon) = q_0^* + q_1(\varepsilon)$, with q_0^* solution of the ε-independent equation (14) and $q_1(\varepsilon)$ vanishing with ε.

The periodicity of an orbit for the full Hamiltonian (7) is given by

$$\begin{aligned}
\hat q(T) - \hat q(0) - \Lambda T &= \int_0^T \nabla_p \left[f_4^{(r,0)} + \sum_{s=1}^r f_2^{(r,s)} \right] ds + \mathcal{O}(|p|^2) + \mathcal{O}(\varepsilon|p|) + \mathcal{O}\big(\varepsilon^{r+1}\big) = 0 \ ,\\
p_1(T) - p_1(0) &= \mathcal{O}(\varepsilon|p|^2) + \mathcal{O}\big(\varepsilon^{r+1}\big) = 0 \ ,\\
p(T) - p(0) &= -\int_0^T \sum_{s=1}^r \nabla_q \left[f_0^{(r,s)} + f_2^{(r,s)} \right] ds + \mathcal{O}(\varepsilon|p|^2) + \mathcal{O}\big(\varepsilon^{r+1}\big) = 0 \ ,
\end{aligned}$$

where the unknown is the initial datum $(\hat q = \hat q(0), \hat p = \hat p(0))$, namely the Cauchy problem. Due to the conservation of the energy, we can eliminate the equation for p_1, divide the $n-1$ actions p by ε and look at $q_1(0)$ as a parameter (the phase along the orbit). The system of $2n-1$ equations in $2n-1$ unknowns $(q(0), p_1(0), p(0))$

$$\begin{aligned}
\hat q(T) - \hat q(0) - \Lambda T &= \int_0^T \nabla_p \left[f_4^{(r,0)} + \sum_{s=1}^r f_2^{(r,s)} \right] ds + \mathcal{O}(|p|^2) + \mathcal{O}(\varepsilon|p|) + \mathcal{O}\big(\varepsilon^{r+1}\big) = 0 \ ,\\
\frac{p(T) - p(0)}{\varepsilon} &= -\frac{1}{\varepsilon} \int_0^T \sum_{s=1}^r \nabla_q \left[f_0^{(r,s)} + f_2^{(r,s)} \right] ds + \mathcal{O}(|p|^2) + \mathcal{O}(\varepsilon^r) = 0 \ ,
\end{aligned}$$

takes the form (10). The approximate periodic solution

$$\hat{p}(t) = 0 \ , \qquad q_1(t) = \omega t + q_1(0) \ , \qquad q(t) = q^* \ ,$$

corresponds to (and actually represents) an approximate zero ($q(0) = q^*, \hat{p}(0) = 0$) for the Υ map. The proof of Theorem 1.1 then simply consists in the application of

Proposition 3.1 (Newton-Kantorovich method) *Consider* $\Upsilon \in \mathcal{C}^1(\mathcal{U}(x_0) \times \mathcal{U}(0), V)$. *Assume that there exist three constants* $C_{1,2,3} > 0$ *dependent, for* ε *small enough, on* $\mathcal{U}(x_0) \subset V$ *only, and two parameters* $0 \le 2\alpha < \beta$ *such that*

$$\begin{aligned} \|\Upsilon(x_0, \varepsilon)\| &\le C_1 |\varepsilon|^\beta \ , \\ \|[\Upsilon'(x_0, \varepsilon)]^{-1}\|_{\mathcal{L}(V)} &\le C_2 |\varepsilon|^{-\alpha} \ , \\ \left\|\Upsilon'(z, \varepsilon) - \Upsilon'(x_0, \varepsilon)\right\|_{\mathcal{L}(V)} &\le C_3 \left\|z - x_0\right\| \ . \end{aligned} \tag{28}$$

Then there exist positive C_0 *and* ε^* *such that, for* $|\varepsilon| < \varepsilon^*$*, there exists a unique* $x^*(\varepsilon) \in \mathcal{U}(x_0)$ *which fulfills*

$$\Upsilon(x^*, \varepsilon) = 0 \ , \qquad \qquad \|x^* - x_0\| \le C_0 |\varepsilon|^{\beta - \alpha} \ .$$

Furthermore, Newton's algorithm converges to x^**.*

The proof of the Proposition is reported in Appendix B. We recall that $\|\cdot\|_{\mathcal{L}(V)}$ represents the usual norm for a linear operator from V to V.

Since we are seeking for a true periodic solution close to the approximate one, we take $(q, \hat{p})$ in a small ball centered in $(q^*, 0)$, that plays the role of x_0 in Proposition 3.1. Thus both the variables can be interpreted "locally" as cartesian variables in $\mathbb{R}^{2n-1}$. We have already introduced $M(\varepsilon)$ in (11), being the differential of the map Υ evaluated in $(q, \hat{p}) = (q^*, 0)$. Extracting from $M(\varepsilon)$ its leading order in ε, we get

$$M(\varepsilon) = M_0 + \tilde{M}_1(\varepsilon) \ , \qquad M_0 := M(0) = \begin{pmatrix} 0 & C_0 \\ B_0 & D_0 \end{pmatrix} \ ,$$

where

$$B_{0;i,j} = -\left[\frac{\partial^2 f_0^{(r,1)}}{\partial q_i \partial q_j}\Big|_{q = q_0^*}\right] \frac{T}{\varepsilon} \ , \qquad \qquad C_0 = CT \ , \tag{29}$$

and C is the twist matrix defined in (4). The first of (28) is satisfied with $\beta = r$. The third of (28) is satisfied in view of the smoothness of the flow at time T w.r.t. the initial datum (it keeps the same smoothness as its vector field). The core of the statement is then the requirement on the invertibility of $M(\varepsilon)$. If B_0 is invertible, then the same holds true for M_0 (the twist being C_0 invertible) which is the leading order of M; hence $M(\varepsilon)$ is also invertible and the second of (28) is satisfied with $\alpha = 0$, M_0 being independent of ε. This is actually Poincaré's theorem. If instead B_0 has a nontrivial Kernel, then the same holds also for M_0, typically with a greater dimension. The required invertibility of $M(\varepsilon)$, asked by Theorem 1.1, is necessarily due to the ε-corrections, which are responsible for the bifurcations of the zero eigenvalues of the matrix M_0. Hence, in order to fulfill the second of (28), we need the smallest eigenvalues of $M(\varepsilon)$ to bifurcate from zero as $\lambda_j(\varepsilon) \sim \varepsilon^\alpha$, with $\alpha < \frac{r}{2}$, which is indeed (12). Finally, estimates (13) are of the same type as the one in Proposition 3.1, even after back-transforming the solutions to the original canonical variables with $\Phi^{(r)}$. Indeed, as illustrated in the detailed proof of Proposition 1.1, the normalizing transformation $\Phi^{(r)}$ is a near the identity transformation.

4 One parameter families.

Generically we expect that, apart from very pathological examples, two normal form steps are enough to get a clear insight into the degeneracy. In particular, with a second order approximation one can investigate whether one-parameter families $q_0^*(s)$, which are solutions of (14), are destroyed or not. In the first case, the isolated solutions which survive to the breaking of the family are natural candidates for continuation, once (12) has been verified. In the second case, at least a third step of normalization is necessary, unless there are good reasons to believe that the whole family survives, due to the effect of some hidden symmetry of the model.

What we are going to develop in the first part of this Section is exactly the case when the first of (29) admits one-parameter families of solutions on the torus $\mathbb{T}^{n-1}$, which means that $\dim(\mathrm{Ker}(B_0)) = 1$. In this easier case (which represents the weakest degeneracy for B_0), under suitable conditions on the matrix M_0, it is possible to apply some results of perturbation theory of matrices to $M(\varepsilon)$ (see [44], Chap. IV, par. 1.4) in order to replace assumption (12) with a more accessible criterion. This allows to get a more applicable formulation of Theorem 1.1, which will be used in the forthcoming application.

4.1 Some few facts on matrix perturbation theory

The degeneration we are considering here implies that $0 \in \mathrm{Spec}(B_0)$, with the geometric multiplicity being equal to one ($m_g(0, B_0) = 1$). Let a_1 be the $(n-1)$-dimensional vector generating $\mathrm{Ker}(B_0)$. Let us introduce also f_1 as the embedding of a_1 into $\mathbb{R}^{2n-1}$, namely the $(2n-1)$ vector

$$f_1 = \begin{pmatrix} a_1 \\ 0 \end{pmatrix} .$$

We have the following

Lemma 4.1 *Assume that the kernel of M_0 had dimension one and is generated by f_1, namely* $\mathrm{Ker}(M_0) = \mathrm{Span}(f_1)$ *. If the orthogonality condition*

$$\left\langle C_0^{-1} D_0^\top a_1, \begin{pmatrix} a_1 \\ 0 \end{pmatrix} \right\rangle = 0 , \tag{30}$$

is fulfilled, then the algebraic multiplicity of the zero eigenvalue is greater than two ($m_a(0, M_0) \geq 2$).

Proof. In order to study the $\mathrm{Ker}(M_0)$, we have to solve

$$\begin{pmatrix} O & C_0 \\ B_0 & D_0 \end{pmatrix} \begin{pmatrix} x \\ -y \end{pmatrix} = \begin{pmatrix} C_0 y \\ B_0 x - D_0 y \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \end{pmatrix}$$

which, due to the invertibility of C_0, gives $y = 0$, and thus $x \in \mathrm{Ker}(B_0)$. This provides the first claim. The statement concerning the algebraic multiplicity can be derived investigating the Kernel of the adjoint matrix $M_0^\top$. It is easy to see that

$$\mathrm{Ker}(M_0^\top) = \mathrm{Span}\,(g) , \qquad g = \begin{pmatrix} -C_0^{-1} D_0^\top a_1 \\ a_1 \end{pmatrix}$$

and to deduce that the assumption (30) is equivalent to $\langle f_1, g\rangle = 0$, where the right hand vector in (30) is the n-dimensional vector built by complementing a_1 with one 0. The last, according to Lemma III, Chapter 1.16 of [44], is not compatible with $m_a(0, M_0) = 1$. Precisely, we can observe that the orthogonality condition between the two vectors allows to find a second generalized eigenvector f_2 for $\mathrm{Ker}(M_0)$, as a solution of $M_0 f_2 = f_1$. Indeed, the Fredholm alternative theorem guarantees the existence of f_2 under exactly the condition $\langle f_1, g\rangle = 0$.

□

In order to determine the asymptotic behavior of the eigenvalues $\lambda(\varepsilon) \in \mathrm{spec}(M(\varepsilon))$, we make use of the fact that $dim(\mathrm{Ker}(M_0)) = 1$ and we assume some minimal smoothness of $M(\varepsilon)$ with respect to ε, namely that the following expansion holds[3]

$$M(\varepsilon) = M_0 + \varepsilon M_1 + \mathcal{O}(\varepsilon^2) = \begin{pmatrix} \varepsilon A_1 & C_0 + \varepsilon C_1 \\ B_0 + \varepsilon B_1 & D_0 + \varepsilon D_1 \end{pmatrix} + \mathcal{O}(\varepsilon^2) .$$

Then the following Lemma holds true (see [44], Chapter IV, § 1, for all the details)

Lemma 4.2 *Let λ_0 be an eigenvalue M_0 with $m_g(\lambda_0, M_0) = 1$ and $m_a(\lambda_0, M_0) = h \geq 2$ and let $f_1, \ldots, f_h$ be the generalized eigenvectors relative to λ_0, defined by the recursive scheme*

$$M_0 f_1 = \lambda_0 f_1, \quad M_0 f_2 = \lambda_0 f_2 + f_1, \ldots, M_0 f_h = \lambda_0 f_h + f_{h-1}.$$

Moreover, let $g_1, \ldots, g_h$ be the generalized eigenvectors for $M_0^\top$ relative to λ_0, such that

$$\langle f_j, g_i\rangle = \delta_{ji}, \quad con \quad j, i = 1, \ldots, h$$

and define

$$\gamma = \langle M_1 f_1, g_h\rangle .$$

If $\gamma \neq 0$, then the h solutions $\lambda_j(\varepsilon)$ of the characteristic equation

$$\det(M(\varepsilon) - \lambda I) = 0$$

are given by

$$\lambda_j(\varepsilon) = \lambda_0 - (\varepsilon\gamma)_j^{1/h} + \mathcal{O}(\varepsilon^{2/h}) ,$$

where $(\varepsilon\gamma)_j^{1/h}$ are the h distinct roots of $\sqrt[h]{\varepsilon\gamma}$.

4.2 The special case of $m_a(0, M_0) = 2$.

We are interested in the bifurcations of the zero eigenvalue (needed to bound the inverse matrix $M^{-1}(\varepsilon)$), thus in the previous Lemma 4.2 we can take $\lambda_0 = 0$ and f_1 as the eigenvector generating $\mathrm{Ker}(M_0)$. Moreover, since

$$\begin{pmatrix} A_1 & C_1 \\ B_1 & D_1 \end{pmatrix} \begin{pmatrix} a_1 \\ 0 \end{pmatrix} = \begin{pmatrix} A_1 a_1 \\ B_1 a_1 \end{pmatrix} ,$$

[3]This is not an obvious fact, since the smoothness of $M(\varepsilon)$ is related to the smoothness of $q^*(\varepsilon)$, solution of the trigonometric system of equations $\nabla_q \left[f_0^{(2,1)} + f_0^{(2,2)} \right] = 0$.

the value of γ does not depend on the whole matrix M_1, but only on the blocks A_1 and B_1. The problem is further simplified when $m_a(0, M_0) = 2$: in this case g_2 coincides with g and γ reduces to

$$\gamma = \langle M_1 f_1, g_2 \rangle = \Big\langle \big(A_1 a_1 \quad B_1 a_1 \big) , \begin{pmatrix} -C_0^{-1} D_0^\top a_1 \\ a_1 \end{pmatrix} \Big\rangle = \Big\langle \big(B_1 - D_0 C_0^{-1} A_1 \big) a_1, a_1 \Big\rangle \ .$$

Thus, under the easier condition

$$\gamma = \langle \big(B_1 - D_0 C_0^{-1} A_1 \big) a_1, a_1 \rangle \neq 0 \ ,$$

Theorem 1.1 can be formulated as

Theorem 4.1 *Consider $\Upsilon = (F, G)$ defined by* (10) *in a neighbourhood of the point $(q^*, 0)$, with $q^*(\varepsilon) \in C^1(\mathcal{U}(0))$ defined by* (9) *and $r = 2$. Let $\dim(\mathrm{Ker}(B_0)) = 1$, a_1 being its generator. Assume also that $m_a(0, M_0) = 2$ and that*

$$\langle \big(B_1 - D_0 C_0^{-1} A_1 \big) a_1, a_1 \rangle \neq 0 \ . \tag{31}$$

Then, there exist positive constants C_0 and ε^* such that, for $|\varepsilon| < \varepsilon^*$ there exists a point $(q_{\mathrm{p.o.}}(\varepsilon), \hat{p}_{\mathrm{p.o.}}(\varepsilon)) \in \mathcal{U} \times \mathbb{T}^{n-1}$ which solves

$$\Upsilon(q_{\mathrm{p.o.}}, \hat{p}_{\mathrm{p.o.}}; \varepsilon, q_1(0)) = 0 \ , \qquad \qquad \|(q_{\mathrm{p.o.}}, \hat{p}_{\mathrm{p.o.}}) - (q^*, 0)\| \leq C_0 \varepsilon^{3/2} \ .$$

In order to verify condition (31), the block matrices A_1 and B_1 are needed; as a consequence, the first order corrections to the generic Cauchy problem, $\hat{q}^{(1)}(t)$ and $\hat{p}^{(1)}(t)$ have to be derived. With a standard approach, as the one performed in [25], and after expanding in ε both the period map Υ and the solution $q^*(\varepsilon) = q_0 + \mathcal{O}(\varepsilon)$ one gets

$$\begin{aligned} \varepsilon A_1 &= -\frac{T^2}{2} C_0 D_q \nabla_{\hat{q}} f_0^{(2,1)}(q_0^*) + T D_q \nabla_{\hat{p}} f_2^{(2,1)}(q_0^*) \\ \varepsilon B_1 &= -T D_q^3 f_0^{(2,1)}(q_0^*) q_1^* - \frac{T}{\varepsilon} D_q^2 f_0^{(2,2)}(q_0^*) \\ &\quad + \frac{T^2}{2} \left[D_{qp}^2 f_2^{(2,1)}(q_0^*) D_q^2 f_0^{(2,1)}(q_0^*) - D_q^2 f_0^{(2,1)}(q_0^*) D_{qp}^2 f_2^{(2,1)}(q_0^*) \right] \\ &\quad + \frac{T^3}{6\varepsilon} \left[D_q^2 f_0^{(2,1)}(q_0^*) C_0 D_q^2 f_0^{(2,1)}(q_0^*) \right] \ . \end{aligned}$$

Despite the formulation of Theorem 4.1 is simplified with respect to the abstract result stated in Theorem 1.1, it is evident from the above formulas that it can be a hard task to verify condition (31). However, if the original Hamiltonian is even in the angle variables, as often happens in models of weakly interacting anharmonic oscillators, then condition (31) can be further simplified if the solutions to be investigated are the in/out-of-phase solutions $q^* = 0, \pi$, as shown in the following example.

4.3 Example: square dNLS cell with nearest neighbour interaction

Let us consider the Hamiltonian system in real coordinates

$$H = H_0 + \varepsilon H_1 = \sum_{j=1}^{4} \left(\frac{x_j^2 + y_j^2}{2} + \left(\frac{x_j^2 + y_j^2}{2} \right)^2 + \varepsilon (x_{j+1} x_j + y_{j+1} y_j) \right) \ ,$$

with periodic boundary conditions, i.e., $x_5 = x_1$ and $y_5 = y_1$. Introducing the action-angle variables $(x_j, y_j) = (\sqrt{2I_j}\cos\varphi_j, \sqrt{2I_j}\sin\varphi_j)$, the Hamiltonian reads

$$H = \sum_{j=1}^{4} \Big(I_j + I_j^2 + 2\varepsilon\sqrt{I_{j+1}I_j}\cos(\varphi_{j+1} - \varphi_j) \Big) \ .$$

Let us now fix the fully resonant torus $I^* = (I^*, I^*, I^*, I^*)$ and make a Taylor expansion around I^*, i.e., we set $I_j = J_j + I^*$ for $j = 1, \ldots, 4$. The unperturbed part, H_0, reads

$$H_0(J) = 4I^* + 4(I^*)^2 + (1 + 2I^*)(J_1 + J_2 + J_3 + J_4) + J_1^2 + J_2^2 + J_3^2 + J_4^2 \ ,$$

while the perturbation H_1 takes the form

$$\begin{aligned} H_1(J,\varphi) = 2I^*(&\cos(\varphi_2 - \varphi_1) + \cos(\varphi_3 - \varphi_2) + \cos(\varphi_4 - \varphi_3) + \cos(\varphi_4 - \varphi_1)) \\ &+ (J_1 + J_2)\cos(\varphi_2 - \varphi_1) + (J_3 + J_2)\cos(\varphi_3 - \varphi_2) \\ &+ (J_4 + J_3)\cos(\varphi_4 - \varphi_3) + (J_1 + J_4)\cos(\varphi_4 - \varphi_1) + \mathcal{O}(|J|^2) \ . \end{aligned}$$

We introduce[4] the resonant angles $\hat{q} = (q_1, q)$ and their conjugate actions $\hat{p} = (p_1, p)$

$$\begin{cases} q_1 = \varphi_1 \\ q_2 = \varphi_2 - \varphi_1 \\ q_3 = \varphi_3 - \varphi_2 \\ q_4 = \varphi_4 - \varphi_3 \end{cases}, \qquad \begin{cases} p_1 = J_1 + J_2 + J_3 + J_4 \\ p_2 = J_2 + J_3 + J_4 \\ p_3 = J_3 + J_4 \\ p_4 = J_4 \end{cases} .$$

Thus, ignoring the constant terms, we can rewrite H as

$$\begin{aligned} H &= \omega p_1 + \Big((p_1 - p_2)^2 + (p_2 - p_3)^2 + (p_3 - p_4)^2 + p_4^2 \Big) + \\ &\quad + \varepsilon \Big[\Big(2I^*\cos(q_2) + 2I^*\cos(q_3) + 2I^*\cos(q_4) + 2I^*\cos(q_2 + q_3 + q_4) \Big) \\ &\qquad + (p_1 - p_3)\cos(q_2) + (p_2 - p_4)\cos(q_3) + p_3\cos(q_4) \\ &\qquad + (p_1 - p_2 + p_4)\cos(q_2 + q_3 + q_4) \Big] + \mathcal{O}(\varepsilon|\hat{p}|^2) \\ &= \omega p_1 + f_4^{(0,0)}(p_1, p_2, p_3, p_4) + f_0^{(0,1)}(q_2, q_3, q_4) \\ &\qquad + f_2^{(0,1)}(p_1, p_2, p_3, p_4, q_2, q_3, q_4) + \mathcal{O}(\varepsilon|\hat{p}|^2) \ , \end{aligned}$$

where $\omega = 1 + 2I^*$.

Remark 4.1 *With the usual canonical complex coordinates* $\psi_j = \frac{1}{\sqrt{2}}(x_j + \mathbf{i}y_j)$*, the Hamiltonian reveals to be a dNLS model, with periodic boundary conditions*

$$H = \sum_{j=1}^{4} \big[|\psi_j|^2 + |\psi_j|^4 + \varepsilon\big(\psi_{j+1}\overline{\psi}_j + c.c.\big) \big] \ , \qquad \psi_5 = \psi_1 \ . \tag{32}$$

In agreement with this, we observe that the Hamiltonian does not depend on the fast angle q_1*. This is due to the effect of the Gauge symmetry of the model, as visible in the complex form (32). As a consequence,* $f_0^{(0,1)}(q_2, q_3, q_4)$ *is already in normal form and the first stage only consists in the translation of the actions, which allows to keep fixed* ω*.*

[4] In this case, we have preferred the angles to be the relative phase differences among consecutive angles, rather than the phase differences with respect to the first angle φ_1.

Since $f_2^{(0,1)}$ is automatically averaged w.r.t. q_1, the homological equation defining $\zeta^{(1)}$ is equivalent to the following linear system

$$\langle \nabla_{\hat{p}} f_4^{(0,0)}, \zeta^{(1)} \rangle = f_2^{(0,1)}\Big|_{q=q^*} ,$$

whose solution is given by

$$\begin{cases} \zeta_1^{(1)} = \varepsilon[\cos(q_2^*) + \cos(q_3^*) + \cos(q_4^*) + \cos(q_2^* + q_3^* + q_4^*)] \\ \zeta_2^{(1)} = \varepsilon\left[\dfrac{\cos(q_2^*)}{2} + \cos(q_3^*) + \cos(q_4^*) + \dfrac{\cos(q_2^* + q_3^* + q_4^*)}{2}\right] \\ \zeta_3^{(1)} = \varepsilon\left[\dfrac{\cos(q_3^*)}{2} + \cos(q_4^*) + \dfrac{\cos(q_2^* + q_3^* + q_4^*)}{2}\right] \\ \zeta_4^{(1)} = \varepsilon\left[\dfrac{\cos(q_4^*)}{2} + \dfrac{\cos(q_2^* + q_3^* + q_4^*)}{2}\right] \end{cases} .$$

Since the normal form preserves the symmetry, the newly generated term $f_2^{(\mathrm{I};0,1)}$ is again independent of q_1 and no further average is required. The values q^*, which define the approximate periodic orbit at leading order, are given by the solutions of the trigonometric system (depending only on sines, due to the parity of the Hamiltonian)

$$\begin{cases} -2I^* \sin(q_2) - 2I^* \sin(q_2 + q_3 + q_4) = 0 \\ -2I^* \sin(q_3) - 2I^* \sin(q_2 + q_3 + q_4) = 0 \\ -2I^* \sin(q_4) - 2I^* \sin(q_2 + q_3 + q_4) = 0 \end{cases} .$$

Such solutions are given by the two isolated configurations $(0,0,0)$, (π,π,π), and the three one-parameter families $Q_1 = (\vartheta, \vartheta, \pi - \vartheta)$, $Q_2 = (\vartheta, \pi - \vartheta, \vartheta)$, $Q_3 = (\vartheta, \pi - \vartheta, \pi - \vartheta)$, with $\theta \in S^1$, which all intersect in the two opposite configurations $\pm(\frac{\pi}{2}, \frac{\pi}{2}, \frac{\pi}{2})$. Since the twist condition (4) is verified, we only need (14) in order to apply the implicit function theorem (which reduces to the classical result of Poincaré). Factoring out $-2I^*$, the nondegeneracy condition reads

$$\left|\begin{pmatrix} \cos(q_2^*) + \cos(q_2^* + q_3^* + q_4^*) & \cos(q_2^* + q_3^* + q_4^*) & \cos(q_2^* + q_3^* + q_4^*) \\ \cos(q_2^* + q_3^* + q_4^*) & \cos(q_3^*) + \cos(q_2^* + q_3^* + q_4^*) & \cos(q_2^* + q_3^* + q_4^*) \\ \cos(q_2^* + q_3^* + q_4^*) & \cos(q_2^* + q_3^* + q_4^*) & \cos(q_4^*) + \cos(q_2^* + q_3^* + q_4^*) \end{pmatrix}\right| \neq 0.$$

If we evaluate the determinant in the two isolated configurations, we get $\det(B_0) = \pm 4T \neq 0$, hence the corresponding solutions can be continued for small enough ε. In the three families we obviously get a degeneration, since the tangent direction to each family represents a Kernel direction, hence $\det\big(B_0\big|_{Q_j}\big) = 0$. Furthermore in the intersections $\pm(\frac{\pi}{2}, \frac{\pi}{2}, \frac{\pi}{2})$ the matrices are identically zero. For all these families a second normalization step is thus needed.

The first stage of the second normalization step deals with

$$f_0^{(1,2)} = f_0^{(\mathrm{I};0,2)} = L_{\langle \zeta^{(1)}, \hat{q} \rangle} f_2^{(0,1)} + \frac{1}{2} L^2_{\langle \zeta^{(1)}, \hat{q} \rangle} f_4^{(0,0)} ,$$

which is already averaged over q_1, due to the preservation of the symmetry. The same holds also for the linear term in the action variables $f_2^{(1,2)}$, given by

$$f_2^{(1,2)} = f_2^{(\mathrm{I};0,2)} = L_{\langle \zeta^{(1)}, \hat{q} \rangle} f_4^{(0,1)} .$$

Hence, the homological equation providing the new translation $\zeta^{(2)}$ reads

$$L_{\langle\zeta^{(2)},\hat{q}\rangle} f_4^{(0,0)} + L_{\langle\zeta^{(1)},\hat{q}\rangle} f_4^{(0,1)}\Big|_{q=q^*} = 0 \ .$$

The new linear term in the action

$$f_2^{(\mathrm{I};1,2)} = L_{\langle\zeta^{(1)},\hat{q}\rangle} f_4^{(0,1)} + L_{\langle\zeta^{(2)},\hat{q}\rangle} f_4^{(0,0)} \ ,$$

is again already averaged over q_1, hence the second step is concluded, and the transformed Hamiltonian reads

$$\begin{aligned} H^{(2)} &= \omega p_1 + f_4^{(2,0)}(\hat{p}) \\ &\quad + f_0^{(2,1)}(q) + f_2^{(2,1)}(\hat{p},q) \\ &\quad + f_0^{(2,2)}(q) + f_2^{(2,2)}(\hat{p},q) \\ &\quad + \mathcal{O}(\varepsilon|\hat{p}|^2) + \mathcal{O}\big(\varepsilon^3\big) \ . \end{aligned}$$

The approximate periodic orbits correspond to the q^* for which

$$\nabla_q\Big[f_0^{(2,1)}(q) + f_0^{(2,2)}(q)\Big] = \nabla_q f_0^{(2,1)}(q) + \nabla_q\Big\langle \nabla_{\hat{p}} f_2^{(0,1)}(q),\, \zeta^{(1)}\Big\rangle = 0 \ ,$$

where in the correction due to $f_0^{(2,2)}$, only the term $L_{\langle\zeta^{(1)},\hat{q}\rangle} f_2^{(0,1)}$ really matters, having a nontrivial dependence on the slow angles q. By exploiting the explicit expression for ζ_1 previously derived, and replacing q^* with q in it, we explicitly get the system

$$\left\{\begin{aligned} &-8\left(\sin(q_2)+\sin(q_2+q_3+q_4)\right)+\varepsilon\Big[2\sin(2q_2)+\sin(q_2-q_3)+2\sin(q_2+q_3) \\ &\qquad + 2\sin(2q_2+2q_3+2q_4)+2\sin(2q_2+q_3+q_4) \\ &\qquad + \sin(q_2+q_3+2q_4)\Big]=0 \\ &-8\left(\sin(q_3)+\sin(q_2+q_3+q_4)\right)+\varepsilon\Big[2\sin(2q_3)+\sin(q_3-q_2)+2\sin(q_2+q_3) \\ &\qquad + \sin(q_3-q_4)+2\sin(q_3+q_4) \\ &\qquad + 2\sin(2q_2+2q_3+2q_4)+\sin(2q_2+q_3+q_4) \\ &\qquad + \sin(q_2+q_3+2q_4)\Big]=0 \\ &-8\left(\sin(q_4)+\sin(q_2+q_3+q_4)\right)+\varepsilon\Big[2\sin(2q_4)+\sin(q_4-q_3)+2\sin(q_3+q_4) \\ &\qquad + 2\sin(2q_2+2q_3+2q_4)+\sin(2q_2+q_3+q_4) \\ &\qquad + 2\sin(q_2+q_3+2q_4)\Big]=0 \end{aligned}\right. ,$$

depending on the effective small parameter $\tilde{\varepsilon} = \frac{\varepsilon}{I^*}$. The above system has the structure

$$F(q,\varepsilon) = F_0(q) + \varepsilon F_1(q) = 0 \ , \tag{33}$$

where $F : \mathbb{T}^3 \times \mathcal{U}(0) \to \mathbb{R}^3$. Moreover, we have already found at first normalization step that

$$F(Q_j(\theta),0) = F_0(Q_j(\theta)) = 0 \ .$$

Suppose that there exists a solution $q(\varepsilon) = (q_2(\varepsilon), q_3(\varepsilon), q_4(\varepsilon))$ which is at least continuous in the small parameter, i.e. $\mathcal{C}^0(\mathcal{U}(0), \mathbb{T}^3)$. Hence, by continuity, we must have

$$\lim_{\varepsilon\to 0} F(q_2(\varepsilon), q_3(\varepsilon), q_4(\varepsilon), \varepsilon) = F_0(q_2(0), q_3(0), q_4(0)) = 0 \ ,$$

which means that $q(0) \in Q_j$. Let us introduce the matrices $\tilde{B}_{0,j}(\vartheta) = \frac{\partial F_0(Q_j(\vartheta))}{\partial q}$ and observe that the tangent directions to the three families

$$\partial_\vartheta Q_1 = \begin{pmatrix} 1 \\ 1 \\ -1 \end{pmatrix}, \qquad \partial_\vartheta Q_2 = \begin{pmatrix} 1 \\ -1 \\ 1 \end{pmatrix} \qquad \text{and} \qquad \partial_\vartheta Q_3 = \begin{pmatrix} 1 \\ -1 \\ -1 \end{pmatrix}$$

represent the Kernel directions of $\tilde{B}_{0,j}$, for $j = 1, 2, 3$, respectively. A standard proposition of bifurcation theory provides a necessary condition for the existence of a solution $Q_j(\theta, \varepsilon)$ which is a continuation of $Q_j(\theta)$.

Proposition 4.1 *Necessary condition for the existence of a solution $q(\varepsilon) = Q_j(\vartheta, \varepsilon)$ of* (33) *is that*

$$F_1(Q_j(\vartheta, 0)) \in \text{Range}(\tilde{B}_{0,j}(\vartheta)) \ .$$

If $\tilde{B}_{0,j}(\vartheta)$ is symmetric, the above condition simplifies

$$F_1(Q_j(\vartheta, 0)) \perp \text{Ker}(\tilde{B}_{0,j}(\vartheta)). \tag{34}$$

Let us apply the above Proposition to show that the families Q_1 and Q_3 break down. Precisely, all their points, except for those corresponding to $\theta = \{0, \pi/2, \pi\}$, do not represent true candidates for the continuation. We compute $\langle F_1(Q_j(\vartheta, 0)), \partial_\theta Q_j\rangle$ for $j = 1, 3$

$$\langle F_1(Q_1(\vartheta)), \partial_\vartheta Q_1\rangle = 8\sin(2\vartheta) = \langle F_1(Q_3(\vartheta)), \partial_\vartheta Q_3\rangle \ ,$$

which shows that the necessary condition is generically violated for the two families $Q_{1,3}$, apart from the in/out-of-phase configurations $(0, 0, \pi)$, $(\pi, \pi, 0)$, $(0, \pi, \pi)$, $(\pi, 0, 0)$ and the symmetric vortex configurations $\pm\left(\frac{\pi}{2}, \frac{\pi}{2}, \frac{\pi}{2}\right)$, the last being also points of $Q_2(\theta)$.

A way to conclude that the above mentioned in/out-of-phase configurations can be continued to periodic solutions is to apply Theorem 4.1. Indeed, the main and first fact to notice is that if $q_0^* = 0, \pi$ then $D_0 = 0$, since it depends only on sines; then by Lemma 4.1 we get $m_a(0, M_0) \geq 2$. Moreover, a direct computation shows that the algebraic multiplicity of the zero eigenvalue of M_0 is exactly two, so that we can apply Theorem 4.1. In order to verify the main condition (31), since $D_0 = 0$, we can restrict to compute only B_1. In the configurations $(0, 0, \pi)$ and $(\pi, \pi, 0)$, we get

$$B_1 = \begin{pmatrix} -2 & -1 & -1 \\ -1 & -2 & -1 \\ -1 & -1 & -2 \end{pmatrix} T + \begin{pmatrix} 16\,(I^*)^2 & 0 & 16\,(I^*)^2 \\ 0 & 32\,(I^*)^2 & 32\,(I^*)^2 \\ 16\,(I^*)^2 & 32\,(I^*)^2 & 48\,(I^*)^2 \end{pmatrix} \frac{T^3}{6} \ ,$$

while, in $(0, \pi, \pi)$ and $(\pi, 0, 0)$, we have

$$B_1 = \begin{pmatrix} -2 & -1 & -1 \\ -1 & -2 & -1 \\ -1 & -1 & -2 \end{pmatrix} T + \begin{pmatrix} 48\,(I^*)^2 & 32\,(I^*)^2 & 16\,(I^*)^2 \\ 32\,(I^*)^2 & 32\,(I^*)^2 & 0 \\ 16\,(I^*)^2 & 0 & 16\,(I^*)^2 \end{pmatrix} \frac{T^3}{6} \ .$$

Anyway, we immediately obtain in all the four cases

$$\gamma = \big\langle \langle B_1, a_1 \rangle, a_1 \big\rangle = -4T \neq 0 \ ,$$

with $a_1 = \partial_\vartheta Q_1$ for the first matrix B_1, and $a_1 = \partial_\vartheta Q_3$ for the second one.

Remark 4.2 *We stress that, for the in/out-of-phase configurations, the true and approximate angles coincide, namely $q_{\mathrm{p.o.}} = q^*$. This is due to the parity of the Hamiltonian in the angles and to the Gauge symmetry; the first implies that the remainder, whatever is its order in ε, only depends on the cosines, hence its p-field vanishes at any combination of 0 and π. The second implies that it does never depend on q_1, p_1 being an exact constant of motion; in other words, the field depends only on slow angles q. In this case, Theorem 4.1 could be simplified.*

It remains to investigate the second family Q_2, which satisfies the necessary condition (34) because it represents a solution for (33), namely $F(Q_2(\theta)) \equiv 0$.

We explicitly constructed the normal form up to order three by using Mathematica™and checked that this family still persists. This led us to conjecture that it represents a true solution of the problem. Indeed, using the complex coordinates as in (32), we can reformulate the continuation of periodic orbits on the completely resonant torus $I = (I^*, I^*, I^*, I^*)$ by using the usual ansatz

$$\psi_j = e^{-\mathbf{i}\omega t}\phi_j \ ,$$

which provides the stationary equation for the amplitudes ϕ_j

$$\lambda\phi_j = 2\phi_j|\phi_j|^2 + \varepsilon(\mathcal{L}\phi)_j \ , \qquad \lambda = \omega - 1 \ , \qquad (\mathcal{L}\phi)_j = \phi_{j+1} + \phi_{j-1} \ .$$

If we further assume that the continued solutions have the same amplitude at all the sites, $|\phi_j| = a$, and the phase-shifts belong to the second family Q_2

$$\phi_j = a e^{\mathbf{i}\varphi_j} \ , \qquad \varphi = (\varphi_1, \varphi_1 + \theta, \varphi_1 + \pi, \varphi_1 + \theta + \pi) \ ,$$

then we realize that for any $\theta \in S^1$ one has

$$\mathcal{L} e^{\mathbf{i}\varphi(\theta)} = 0 \ .$$

Hence the stationary equation becomes

$$\lambda = 2a^2 = 2I^* \ ,$$

which implies that a two-dimensional resonant torus, embedded in the original unperturbed four dimensional torus, survives for any given ε.

Remark 4.3 *The above formulation would provide a much simpler proof for the existence of the in/out-of-phase periodic orbits for $\varepsilon \neq 0$, by restricting to study the real ϕ configurations solving the stationary equations [18, 19]. However, the role of this example in the present paper is to show how the formal algorithm works and what kind of insights can lead to in investigating the breaking of completely resonant tori.*

5 Conclusions

Motivated by the aim of investigating the continuation of periodic orbits on a completely resonant torus with respect to a small parameter, we have built up an original normal form algorithm for a classical Hamiltonian model of the form (1). This method naturally extends the averaging procedure of Poincaré, which applies only to nondegenerate approximated solutions. Hence, it allows to deal with all those cases when the extrema of the averaged Hamiltonian are not isolated, like the one-parameter families explored in Section 4. The present formulation of the result deals with the case of a maximal torus, hence it is more suitable for applications for few-bodies problems, e.g., in Celestial Mechanics. In this field, the normal form construction proposed here, which provides a highly accurate approximate dynamics, could be effectively implemented with the aid of an algebraic manipulator (see, e.g., [11]). Besides, the use of numerical tools could also include the analysis of the spectrum of $M(\varepsilon)$, which can be approximated at leading order by the Floquet exponents of the approximate periodic orbits. Hence, hypothesis (12) can be verified numerically, by tracking the dependence of the approximate Floquet spectrum on ε in a neighbourhood of the origin.

The normal form algorithm here developed, if suitably extended to completely resonant low-dimensional tori, could also allow to deal with degenerate scenarios which emerge studying discrete solitons in 1D nonlocal discrete nonlinear Schroedinger lattices (like zigzag dNLS, see [32]): in these models, one parameter families of solutions of the averaged Hamiltonian appear when in the model long range interactions (like next-to-nearest neighbourhood) are added. More naturally, one parameter families of approximate solutions, like the ones observed in the application developed in Section 4, appear in the investigation of vortexes in 2D square lattices [30]. In these problems, the only approach which has been till now explored and applied is based on bifurcation methods [30,33] suitably combined with a perturbation scheme. Hence, a different and completely constructive approach would be desirable, especially in terms of possible applications to the above mentioned lattice models with the help of a manipulator. This further and not trivial extension will be worked out in a future publication.

A Technicalities: normal form construction

The appendix is devoted to technical details and proofs related to the normal form construction which have been moved here in order to avoid the overloading of the text.

A.1 Estimates for the $\nu_{r,s}$ sequence

Lemma A.1 *The sequence* $\{\nu_{r,s}\}_{r\geq 0\,,\,s\geq 0}$ *defined in* (25) *is bounded by the exponential growth*

$$\nu_{r,s} \leq \nu_{s,s} \leq \frac{100^s}{20} \qquad \textit{for} \quad r\geq 0\,,\ s\geq 0\ .$$

Proof. We start with the elimination of $\nu_{r,s}^{(\mathrm{I})}$ in the definition of $\nu_{r,s}$

$$\begin{aligned}\nu_{r,s} &= \sum_{j=0}^{\lfloor s/r\rfloor}(3\nu_{r-1,r})^j \sum_{i=0}^{\lfloor s/r\rfloor-j}(\nu_{r-1,r})^i\nu_{r-1,s-(i+j)r}\\ &= \sum_{j=0}^{\lfloor s/r\rfloor}(3\nu_{r-1,r})^j \sum_{i=j}^{\lfloor s/r\rfloor}(\nu_{r-1,r})^{i-j}\nu_{r-1,s-ir}\\ &= \sum_{i=0}^{\lfloor s/r\rfloor}(\nu_{r-1,r})^i\nu_{r-1,s-ir}\sum_{j=0}^{i}3^j = \sum_{i=0}^{\lfloor s/r\rfloor}\frac{3^{i+1}-1}{2}(\nu_{r-1,r})^i\nu_{r-1,s-ir}\ ,\end{aligned}$$

where in the second equality we have exploited $\nu_{r,0}^{(\mathrm{I})} = 1$. Thus we can rewrite the sequence as

$$\nu_{r,s} = \sum_{j=0}^{\lfloor s/r\rfloor}\theta_j\nu_{r-1,r}^j\nu_{r-1,s-jr}\ , \qquad \theta_j = \frac{3^{j+1}-1}{2}\ .$$

It is immediate to notice that $\nu_{r,s} \le \nu_{s,s}$ for $s \ge r$, hence

$$\nu_{0,s} \le \nu_{1,s} \le \ldots \le \nu_{s,s} = \nu_{s+1,s} = \ldots\ .$$

Moreover

$$\theta_0 = 1\ , \qquad \theta_1 = 4\ , \qquad \theta_{j+1} \le 5\theta_j \qquad \text{for } j \ge 0\ . \tag{35}$$

and observing that $\nu_{r,r} = \theta_0\nu_{r-1,r} + \theta_1\nu_{r-1,r}$, we get

$$\nu_{r,r} = 5\nu_{r-1,r} \qquad \text{for } r \ge 1\ . \tag{36}$$

From the definition of $\{\nu_{r,s}\}$, we can derive the following: for $r \ge 2$ and $s > 2r$ we have

$$\begin{aligned}\nu_{r,s} &= \nu_{r-1,s} + \nu_{r-1,r}\sum_{j=0}^{\lfloor s/r\rfloor-1}\theta_{j+1}\nu_{r-1,r}^j\nu_{r-1,s-r-jr}\\ &\le \nu_{r-1,s} + 5\nu_{r-1,r}\sum_{j=0}^{\lfloor s/r\rfloor-1}\theta_j\nu_{r-1,r}^j\nu_{r-1,s-r-jr}\\ &\le \nu_{r-1,s} + 5\nu_{r-1,r}\nu_{r,s-r} \le \nu_{r-1,s} + \nu_{r,r}\nu_{s-r,s-r}\ ,\end{aligned}$$

where (35) and (36) have been used; for $r = 1$ we have

$$\begin{aligned}\nu_{1,s} &= \nu_{0,s} + \nu_{0,1}\sum_{j=0}^{s-1}\theta_{j+1}\nu_{0,1}^j\nu_{0,s-1-j}\\ &\le (1+\theta_1)\nu_{0,s-1} + 5\sum_{j=1}^{s-1}\theta_j\nu_{0,1}^j\nu_{0,s-1-j}\\ &\le 5\nu_{1,s-1} \le 5\nu_{s-2,s-1} = \nu_{s-1,s-1}\ ,\end{aligned}$$

where (35) has been used, together with $\nu_{0,s} = 1$, for $s \geq 0$. Due to the above properties, we can estimate $\{\nu_{r,s}\}_{r\geq 0\,,\, s\geq 0}$ by means of its diagonal terms $\nu_{r,r}$. Indeed, $\nu_{1,1} = 5$ and for $s > 2$

$$\begin{aligned}\nu_{r,r} &= 5\nu_{r-1,r} \leq 5\nu_{r-2,r} + 5\nu_{r-1,r-1}\nu_{1,1} \leq \ldots \\ &\leq 5\nu_{1,r} + 5\left(\nu_{2,2}\nu_{r-2,r-2} + \ldots + \nu_{r-1,r-1}\nu_{1,1}\right) \leq 5\sum_{j=1}^{r-1} \nu_{j,j}\nu_{r-j,r-j}\ .\end{aligned}$$

From this last upper bound, it is possible to verify

$$\nu_{r,r} \leq 5^{2r-1}\lambda_r \qquad \text{for } r \geq 1\ ,$$

with $\{\lambda_r\}_{r\geq 1}$ being the Catalan sequence, which satisfies $\lambda_r \leq 4^{r-1}$, thus

$$\nu_{r,s} \leq \nu_{s,s} \leq \frac{100^s}{20} \qquad \text{for} \quad r \geq 0\,,\ s \geq 0\ .$$

□

A.2 Estimates for multiple Poisson brackets

Some Cauchy estimates on the derivatives in the restricted domains will be useful.

Lemma A.2 *Let $d \in \mathbb{R}$ such that $0 < d < 1$ and $g \in \mathcal{P}_{2l}$ be an analytic function with bounded norm $\|g\|_1$. Then one has*

$$\left\|\frac{\partial g}{\partial \hat{p}_j}\right\|_{1-d} \leq \frac{\|g\|_1}{d\rho}\,, \qquad \left\|\frac{\partial g}{\partial \hat{q}_j}\right\|_{1-d} \leq \frac{\|g\|_1}{ed\sigma}\ ,$$

Proof. Given g as in (3), one has

$$\begin{aligned}\left\|\frac{\partial g}{\partial \hat{p}_j}\right\|_{1-d} &\leq \sum_{\substack{i\in\mathbb{N}^n\\ |i|=l}} \sum_{k\in\mathbb{Z}^n} \frac{i_j}{\rho}|g_{i,k}|(1-d)^{l-1}\rho^l e^{|k|(1-d)\sigma} \\ &\leq \frac{1}{d\rho}\sum_{\substack{i\in\mathbb{N}^n\\ |i|=l}} \sum_{k\in\mathbb{Z}^n} |g_{i,k}|\rho^l e^{|k|\sigma} = \frac{\|g\|_1}{d\rho}\ ,\end{aligned}$$

where we have used the elementary inequality $m(\lambda - x)^{m-1} \leq \lambda^m/x$, for $0 < x < \lambda$ and $m \geq 1$.

Similarly,

$$\begin{aligned}\left\|\frac{\partial g}{\partial \hat{q}_j}\right\|_{1-d} &\leq \sum_{\substack{i\in\mathbb{N}^n\\ |i|=l}} \sum_{k\in\mathbb{Z}^n} |k_j|\,|g_{i,k}|(1-d)^l\rho^l e^{|k|(1-d)\sigma} \\ &\leq \frac{1}{ed\sigma}\sum_{\substack{i\in\mathbb{N}^n\\ |i|=l}} \sum_{k\in\mathbb{Z}^n} |g_{i,k}|\rho^l e^{|k|\sigma} = \frac{\|g\|_1}{ed\sigma}\ ,\end{aligned}$$

in view of the elementary inequality $x^\alpha e^{-\delta x} \leq (\alpha/(e\delta))^\alpha$, for positive α, x and δ.

□

Lemma A.3 *Let $d \in \mathbb{R}$ such that $0 < d < 1$. Let the generating functions $\chi_0^{(r)}$ and $\chi_2^{(r)}$ be as in* (15)*. Then one has*

$$\left\| \frac{\partial \chi_0^{(r)}}{\partial \hat{q}_j} \right\|_{1-d} \le \frac{\|X_0^{(r)}\|_1}{ed\sigma} + |\zeta^{(r)}| \ , \tag{37}$$

$$\left\| \frac{\partial \chi_2^{(r)}}{\partial \hat{q}_j} \right\|_{1-d} \le \frac{\|\chi_2^{(r)}\|_1}{ed\sigma} \ , \tag{38}$$

$$\left\| \frac{\partial \chi_2^{(r)}}{\partial \hat{p}_j} \right\|_{1-d} \le \frac{\|\chi_2^{(r)}\|_1}{d\rho} \ ; \tag{39}$$

moreover, for $j \ge 1$,

$$\left\| L^j_{\chi_0^{(r)}} f \right\|_{1-d-d'} \le \frac{j!}{e} \left(\frac{\|X_0^{(r)}\|_{1-d'}}{d^2 \rho\sigma} + \frac{e|\zeta^{(r)}|}{d\rho} \right)^j \|f\|_{1-d'} \ , \tag{40}$$

$$\left\| L^j_{\chi_2^{(r)}} f \right\|_{1-d-d'} \le \frac{j!}{e} \left(\frac{\|\chi_2^{(r)}\|_{1-d'}}{d^2 \rho\sigma} \right)^j \|f\|_{1-d'} , \tag{41}$$

Proof. The proofs of (37)–(39) are just minor modifications of Lemma A.2, thus they are left to the reader.

Coming to (40), let $\delta = d/j$ with $j \ge 1$. Proceeding iteratively we get

$$\begin{aligned} \left\| L^j_{\chi_0^{(r)}} f \right\|_{1-d-d'} &\le \left(\frac{\|X_0^{(r)}\|_{1-d'}}{j\delta^2 e\rho\sigma} + \frac{|\zeta^{(r)}|}{\delta\rho} \right) \left\| L^{j-1}_{\chi_0^{(r)}} f \right\|_{1-d'-(j-1)\delta} \\ &\le \ldots \\ &\le \frac{j!}{e} \left(\frac{\|X_0^{(r)}\|_{1-d'}}{d^2\rho\sigma} + \frac{e|\zeta^{(r)}|}{d\rho} \right)^j \|f\|_{1-d'} \ , \end{aligned}$$

where we have used the trivial inequality $j^j \le j!\, e^{j-1}$, holding true for $j \ge 1$. Finally, the proof of (41) is the same, mutatis mutandis.

□

A.3 Estimates for the generating functions

Lemma A.4 *Let $d \in \mathbb{R}$ such that $0 < d < 1$. The generating function $X_0^{(r)}$ and the vector $\zeta^{(r)}$ are bounded by*

$$\|X_0^{(r)}\|_{1-d} \le \frac{\|f_0^{(r-1,r)}\|_{1-d}}{\omega} \ , \qquad |\zeta^{(r)}| \le \frac{\|f_2^{(r-1,r)}\|_{1-d}}{m\rho} \ . \tag{42}$$

The generating function $\chi_2^{(r)}$ is instead bounded by

$$\|\chi_2^{(r)}\|_{1-d} \le \frac{1}{\omega} \left(2\|f_2^{(r-1,r)}\|_{1-d} + \frac{2}{e\delta_r\rho\sigma} \frac{\|f_0^{(r-1,r)}\|_{1-d}}{\omega} \|f_4^{(0,0)}\|_1 \right) \ . \tag{43}$$

Proof. The estimate for $X_0^{(r)}$ is trivial. The estimate for $\chi_2^{(r)}$, that is controlled by $f_2^{(\mathrm{I};r-1,r)}$, is a little bit tricky. Indeed, one has to explicitly exploit the fact that

$$f_2^{(\mathrm{I};r-1,r)} = f_2^{(r-1,r)} - \langle f_2^{(r-1,r)}(q^*)\rangle_{q_1} + L_{X_0^{(r)}} f_4^{(0,0)} \ ,$$

together with the trivial estimate

$$\|f - \langle f(q^*)\rangle_{q_1}\|_{1-d} \leq 2\|f\|_{1-d} \ .$$

As C satisfies (4), there exists a solution $\zeta^{(r)}$ of (18) which satisfies

$$\left\|\nabla_{\hat{p}}\big\langle f_2^{(r-1,r)}\big|_{q=q^*}\big\rangle_{q_1}\right\|_{1-d_{r-1}} = \Big|\sum_j C_{ij}\zeta_j^{(r)}\Big| \geq m|\zeta^{(r)}| \ .$$

Moreover, by the definition of the norm one has

$$\left\|\nabla_{\hat{p}}\big\langle f_2^{(r-1,r)}\big|_{q=q^*}\big\rangle_{q_1}\right\|_{1-d_{r-1}} = \frac{\left\|\big\langle f_2^{(r-1,r)}\big|_{q=q^*}\big\rangle_{q_1}\right\|_{1-d_{r-1}}}{\rho} \leq \frac{\left\|f_2^{(r-1,r)}\right\|_{1-d_{r-1}}}{\rho} \ .$$

Combining the latter inequalities one gets (42).

□

A.4 Estimates for the first and second normalization step

The following two Lemmas collect the estimates concerning the first two steps of the normal form algorithm previously described. We decide to explicitly report the results concerning the normal form at order one and two with the purpose of making transparent the structure of the estimates of the different terms appearing in the normalized Hamiltonian. Furthermore, the first two steps are needed so as to verify the inductive proof for the forthcoming Lemma 2.2.

Lemma A.5 *Consider a Hamiltonian $H^{(0)}$ expanded as in* (16)*. Let $\chi_0^{(1)}$ and $\chi_2^{(1)}$ be the generating functions used to put the Hamiltonian in normal form at order one, then one has*

$$\|X_0^{(1)}\|_1 \leq \frac{1}{\omega}\nu_{0,1}E\varepsilon \ ,$$
$$|\zeta^{(1)}| \leq \frac{1}{4m\rho}\nu_{0,1}E\varepsilon \ ,$$
$$\|\chi_2^{(1)}\|_{1-\delta_1} \leq \frac{1}{\omega}3\nu_{0,1}\Xi_1\frac{E}{4}\varepsilon \ .$$

The terms appearing in the expansion of $H^{(\mathrm{I};0)}$, i.e. in (19) *with $r=1$, are bounded as*

$$\|f_0^{(\mathrm{I};0,1)}\|_{1-\delta_1} \leq E\varepsilon \ ,$$
$$\|f_{2l}^{(\mathrm{I};0,s)}\|_{1-\delta_1} \leq \nu_{1,s}^{(\mathrm{I})}\Xi_1^s\frac{E}{2^l}\varepsilon^s \ .$$

The terms appearing in the expansion of $H^{(1)}$, i.e. in (21) *with $r=1$, are bounded as*

$$\begin{aligned} \|f_0^{(1,s)}\|_{1-d_1} &\leq \nu_{1,s}\Xi_1^{2s-2}E\varepsilon^s \ , \\ \|f_2^{(1,s)}\|_{1-d_1} &\leq \nu_{1,s}\Xi_1^{2s-1}\frac{E}{2^2}\varepsilon^s \ , \\ \|f_{2l}^{(1,s)}\|_{1-d_1} &\leq \nu_{1,s}\Xi_1^{2s}\frac{E}{2^{2l}}\varepsilon^s \ . \end{aligned} \tag{44}$$

Proof. Using Lemma A.4, we immediately get the bounds

$$\left\|X_0^{(1)}\right\|_1 \le \frac{1}{\omega}\left\|f_0^{(0,1)}\right\|_1 \le \frac{1}{\omega}E\varepsilon\ , \qquad |\zeta^{(1)}| \le \frac{1}{m\rho}\left\|f_2^{(0,1)}\right\|_1 \le \frac{E\varepsilon}{4m\rho}\ ,$$

thus, from (37) with $r=1$ we get

$$\left\|\frac{\partial\chi_0^{(1)}}{\partial\hat{q}_j}\right\|_{1-\delta_1} \le \frac{E\varepsilon}{\omega e\delta_1\sigma} + \frac{E\varepsilon}{4m\rho} \le \left(\frac{1}{\omega e\delta_1\sigma} + \frac{1}{4m\rho}\right)E\varepsilon\ .$$

The terms $f_{2l}^{(\mathrm{I};0,s)}$ appearing in the expansion of the Hamiltonian $H^{(\mathrm{I};0)}$ are bounded as follows. For $l=0$ and $s=1$ one has

$$\|f_0^{(\mathrm{I};0,1)}\|_{1-\delta_1} \le \|f_0^{(0,1)}\|_{1-\delta_1} \le E\varepsilon\ , \tag{45}$$

while for the remaining terms one has

$$\begin{aligned}
\|f_{2l}^{(\mathrm{I};0,s)}\|_{1-\delta_1} &\le \sum_{j=0}^{s}\frac{1}{j!}\|L^j_{\chi_0^{(1)}}f_{2l+2j}^{(0,s-j)}\|_{1-\delta_1} \\
&\le \sum_{j=0}^{s}\frac{1}{e}\left(\frac{\|X_0^{(1)}\|_{1-d}}{\delta_1^2\rho\sigma} + \frac{e|\zeta^{(1)}|}{\delta_1\rho}\right)^j\|f_{2l+2j}^{(0,s-j)}\|_1 \\
&\le \sum_{j=0}^{s}\frac{1}{e}\left(\frac{1}{\omega\delta_1^2\rho\sigma} + \frac{e}{4m\delta_1\rho^2}\right)^j E^j\varepsilon^j\frac{E}{2^{2l+2j}}\varepsilon^{s-j} \\
&\le \frac{E\varepsilon^s}{2^{2l}}\sum_{j=0}^{s}\frac{1}{e}\left(\frac{E}{\omega\delta_1^2\rho\sigma} + \frac{eE}{4m\delta_1\rho^2}\right)^j \\
&< (s+1)\Xi_1^s\frac{E}{2^{2l}}\varepsilon^s = \nu_{1,s}^{(\mathrm{I})}\Xi_1^s\frac{E}{2^{2l}}\varepsilon^s,
\end{aligned}$$

where we used the definition of the constant Ξ_1 and Lemma A.3.

Coming to the second stage of the normalization step, the generating function $\chi_2^{(1)}$ is bounded, as in (43), by

$$\begin{aligned}
\|\chi_2^{(1)}\|_{1-\delta_1} &\le \frac{1}{\omega}\left(2\|f_2^{(0,1)}\|_1 + \frac{2}{e\delta_1\rho\sigma}\frac{\|f_0^{(0,1)}\|_{1-\delta_1}}{\omega}\|f_4^{(0,0)}\|_1\right) \\
&\le \frac{1}{\omega}\left(2\frac{E}{4}\varepsilon + \frac{2}{e\delta_1\rho\sigma}\frac{E\varepsilon}{\omega}\frac{E}{2^4}\right) \\
&\le \frac{1}{\omega}\left(2 + \frac{E}{2\omega e\delta_1\rho\sigma}\right)\frac{E}{4}\varepsilon \\
&< \frac{1}{\omega}3\nu_{0,1}\Xi_1\frac{E}{4}\varepsilon\ .
\end{aligned}$$

The terms $f_{2l}^{(1,s)}$ appearing in the expansion of the Hamiltonian $H^{(1)}$ are bounded as follows. The term $f_0^{(1,1)}$ is unchanged, while for $l=0$ and $s=2$ one has

$$\begin{aligned}
\|f_0^{(1,2)}\|_{1-d_1} &\le \|f_0^{(\mathrm{I};0,2)}\|_{1-\delta_1} + \frac{1}{e}\frac{1}{\delta_1^2\rho\sigma}\|\chi_2^{(1)}\|_{1-\delta_1}\|f_0^{(\mathrm{I};0,1)}\|_{1-\delta_1} \\
&\le \nu_{1,2}^{(\mathrm{I})}\Xi_1^2E\varepsilon^2 + \frac{1}{e}\frac{1}{\delta_1^2\rho\sigma}\frac{1}{\omega}3\nu_{0,1}\Xi_1\frac{E}{4}\varepsilon E\varepsilon \\
&\le \nu_{1,2}\Xi_1^2E\varepsilon^2\ .
\end{aligned}$$

For $l=0$ and $s\geq 3$, using (45) for the estimate of the last term in the sum, one has

$$\begin{aligned}\|f_0^{(1,s)}\|_{1-d_1} &\leq \sum_{j=0}^{s-2}\frac{1}{e}\left(\frac{1}{\delta_1^2\rho\sigma}\right)^j\|\chi_2^{(1)}\|_{1-\delta_1}^j\|f_0^{(\mathrm{I};0,s-j)}\|_{1-\delta_1}\\ &\quad+\frac{1}{e}\left(\frac{1}{\delta_1^2\rho\sigma}\right)^{s-1}\|\chi_2^{(1)}\|_{1-\delta_1}^{s-1}\|f_0^{(\mathrm{I};0,1)}\|_{1-\delta_1}\\ &\leq \sum_{j=0}^{s-2}\frac{1}{e}\left(\frac{1}{\delta_1^2\rho\sigma}\right)^j\frac{1}{\omega^j}(3\nu_{0,1})^j\Xi_1^j\frac{E^j}{4^j}\varepsilon^j\nu_{1,s-j}^{(\mathrm{I})}\Xi_1^{s-j}E\varepsilon^{s-j}\\ &\quad+\frac{1}{e}\left(\frac{1}{\delta_1^2\rho\sigma}\right)^{s-1}\frac{1}{\omega^{s-1}}(3\nu_{0,1})^{s-1}\Xi_1^{s-1}\frac{E^{s-1}}{4^{s-1}}\varepsilon^{s-1}E\varepsilon\\ &\leq \nu_{1,s}\Xi_1^{2s-2}E\varepsilon^s\ .\end{aligned}$$

The term $f_2^{(1,1)}$ is unchanged, while for $l=1$ and $s\geq 2$ one has

$$\begin{aligned}\|f_2^{(1,s)}\|_{1-d_1} &\leq \sum_{j=0}^{s-2}\frac{1}{e}\left(\frac{1}{\delta_1^2\rho\sigma}\right)^j\|\chi_2^{(1)}\|_{1-\delta_1}^j\|f_2^{(\mathrm{I};0,s-j)}\|_{1-\delta_1}\\ &\quad+\frac{1}{e}\left(\frac{1}{\delta_1^2\rho\sigma}\right)^{s-1}\|\chi_2^{(1)}\|_{1-\delta_1}^{s-1}\|f_2^{(\mathrm{I};0,1)}\|_{1-\delta_1}+\\ &\leq \sum_{j=0}^{s-2}\frac{1}{e}\left(\frac{1}{\delta_1^2\rho\sigma}\right)^j\frac{1}{\omega^j}(3\nu_{0,1})^j\Xi_1^j\frac{E^j}{4^j}\varepsilon^j\nu_{1,s-j}^{(\mathrm{I})}\Xi_1^{s-j}\frac{E}{4}\varepsilon^{s-j}\\ &\quad+\frac{1}{e}\left(\frac{1}{\delta_1^2\rho\sigma}\right)^{s-1}\frac{1}{\omega^{s-1}}(3\nu_{0,1})^{s-1}\Xi_1^{s-1}\frac{E^{s-1}}{4^{s-1}}\varepsilon^{s-1}\nu_{1,1}^{(\mathrm{I})}\Xi_1\frac{E}{4}\varepsilon\\ &\leq \nu_{1,s}\Xi_1^{2s-1}\frac{E}{2^2}\varepsilon^s\ .\end{aligned}$$

Finally, for $l\geq 2$ and $s\geq 1$ one has

$$\begin{aligned}\|f_{2l}^{(1,s)}\|_{1-d_1} &\leq \sum_{j=0}^{s}\frac{1}{e}\left(\frac{1}{\delta_1^2\rho\sigma}\right)^j\|\chi_2^{(1)}\|_{1-\delta_1}^j\|f_{2l}^{(\mathrm{I};0,s-j)}\|_{1-\delta_1}\\ &\leq \sum_{j=0}^{s}\frac{1}{e}\left(\frac{1}{\delta_1^2\rho\sigma}\right)^j\frac{1}{\omega^j}(3\nu_{0,1})^j\Xi_1^j\frac{E^j}{4^j}\varepsilon^j\nu_{1,s-j}^{(\mathrm{I})}\Xi_1^{(s-j)}\frac{E}{2^{2l}}\varepsilon^{s-j}\\ &\leq \nu_{1,s}\Xi_1^{2s}\frac{E}{2^{2l}}\varepsilon^s\ .\end{aligned}$$

This concludes the proof of the Lemma.

□

Lemma A.6 *Consider a Hamiltonian $H^{(1)}$ expanded as in* (17)*. Let $\chi_0^{(2)}$ and $\chi_2^{(2)}$ be the generating functions used to put the Hamiltonian in normal form at order two, then one has*

$$\begin{aligned}\|X_0^{(2)}\|_{1-d_1} &\leq \frac{1}{\omega}\nu_{1,2}\Xi_2^2E\varepsilon^2\ ,\\ |\zeta^{(2)}| &\leq \frac{1}{4m\rho}\nu_{1,2}\Xi_2^3E\varepsilon^2\ ,\\ \|\chi_2^{(2)}\|_{1-d_1-\delta_2} &\leq \frac{1}{\omega}3\nu_{1,2}\Xi_2^3\frac{E}{4}\varepsilon^2\ .\end{aligned}$$

The terms appearing in the expansion of $H^{(\mathrm{I};1)}$*, i.e. in* (19) *with* $r=2$*, are bounded as*

$$
\begin{aligned}
&\|f_0^{(\mathrm{I};1,s)}\|_{1-d_1-\delta_2} \le \nu_{2,s}^{(\mathrm{I})}\Xi_2^{2s-2}E\varepsilon^s\ , && \textit{for } 1\le s\le 2\ ,\\
&\|f_0^{(\mathrm{I};1,s)}\|_{1-d_1-\delta_2} \le \nu_{2,s}^{(\mathrm{I})}\Xi_2^{2s-1}E\varepsilon^s\ , && \textit{for } 2< s\le 4\ ,\\
&\|f_2^{(\mathrm{I};1,s)}\|_{1-d_1-\delta_2} \le \nu_{2,s}^{(\mathrm{I})}\Xi_2^{2s-1}\frac{E}{2}\varepsilon^s\ , && \textit{for } 1\le s\le 2\ ,\\
&\|f_{2l}^{(\mathrm{I};1,s)}\|_{1-d_1-\delta_2} \le \nu_{2,s}^{(\mathrm{I})}\Xi_2^{2s}\frac{E}{2^{2l}}\varepsilon^s\ , && \textit{for the remaining cases}\ .
\end{aligned}
$$

The terms appearing in the expansion of $H^{(2)}$*, i.e. in* (21) *with* $r=2$*, are bounded as*

$$
\begin{aligned}
&\|f_0^{(2,s)}\|_{1-d_2} \le \nu_{2,s}\Xi_2^{2s-2}E\varepsilon^s\ , && \textit{for } 1\le s\le 2\ ,\\
&\|f_0^{(2,s)}\|_{1-d_2} \le \nu_{2,s}\Xi_2^{2s-1}E\varepsilon^s\ , && \textit{for } 2< s\le 4\ ,\\
&\|f_2^{(2,s)}\|_{1-d_2} \le \nu_{2,s}\Xi_2^{2s-1}\frac{E}{2}\varepsilon^s\ , && \textit{for } 1\le s\le 2\ ,\\
&\|f_{2l}^{(2,s)}\|_{1-d_2} \le \nu_{2,s}\Xi_2^{2s}\frac{E}{2^{2l}}\varepsilon^s\ , && \textit{for the remaining cases}\ .
\end{aligned}
$$

Proof. Using Lemma A.4 and the estimates in Lemma A.5, we immediately get

$$\left\|X_0^{(2)}\right\|_{1-d_1} \le \frac{1}{\omega}\nu_{1,2}\Xi_2^2E\varepsilon^2\ ,\qquad |\zeta^{(2)}| \le \frac{1}{m\rho}\nu_{1,2}\Xi_2^3\frac{E}{4}\varepsilon^2\ ,$$

thus, from (37) we get

$$\left\|\frac{\partial\chi_0^{(2)}}{\partial\hat{q}_j}\right\|_{1-d_1-\delta_2} \le \frac{1}{\omega e\delta_2\sigma}\nu_{1,2}\Xi_2^2E\varepsilon^2 + \frac{1}{4m\rho}\nu_{1,2}\Xi_2^3E\varepsilon^2 \le \left(\frac{1}{\omega e\delta_2\sigma}+\frac{1}{4m\rho}\right)\nu_{1,2}\Xi_2^3E\varepsilon^2\ ,$$

The terms $f_{2l}^{(\mathrm{I};1,s)}$ appearing in the expansion of the Hamiltonian $H^{(\mathrm{I};1)}$ are bounded as follows. For $s=1$ all the terms are unchanged, thus there is nothing to do. Furthermore, notice that $f_0^{(\mathrm{I};1,2)}$ is trivially bounded with the norm of $f_0^{(1,2)}$. The term $f_2^{(\mathrm{I};1,2)}$ requires more care, indeed

$$f_2^{(\mathrm{I};1,2)} = f_2^{(1,2)} - \langle f_2^{(1,2)}(q^*)\rangle_{q_1} + L_{X_0^{(2)}}f_4^{(0,0)}\ .$$

Thus only the generating function $X_0^{(2)}$ plays a role and we get the following estimate

$$\|f_2^{(\mathrm{I};1,2)}\|_{1-d_1-\delta_2} \le 2\nu_{1,2}\Xi_2^3\frac{E}{4}\varepsilon^2 + \frac{1}{\omega e\delta_2\rho\sigma}\nu_{1,2}\Xi_2^2E\varepsilon^2\frac{E}{4} \le 3\nu_{1,2}\Xi_2^3\frac{E}{4}\varepsilon^2 < \nu_{2,2}^{(\mathrm{I})}\Xi_2^3\frac{E}{2}\varepsilon^2\ .$$

For $l=0$ and $s=3$ one has

$$\|f_0^{(\mathrm{I};1,3)}\|_{1-d_1-\delta_2} \le \|f_0^{(1,3)}\|_{1-d_1-\delta_2} + \|L_{\chi_0^{(2)}}f_2^{(1,1)}\|_{1-d_1-\delta_2} \le \nu_{2,3}^{(\mathrm{I})}\Xi_2^4E\varepsilon^3$$

Similarly, for $l=0$ and $s=4$ one has

$$\|f_0^{(\mathrm{I};1,4)}\|_{1-d_1-\delta_2} \le \|f_0^{(1,4)}\|_{1-d_1-\delta_2} + \|L_{\chi_0^{(2)}}f_2^{(1,2)}\|_{1-d_1-\delta_2} + \|L^2_{\chi_0^{(2)}}f_4^{(1,0)}\|_{1-d_1-\delta_2} \le \nu_{2,4}^{(\mathrm{I})}\Xi_2^7E\varepsilon^4\ .$$

For the remaining terms one has

$$\begin{aligned}
\|f_{2l}^{(\mathrm{I};1,s)}\|_{1-d_1-\delta_2} &\le \sum_{j=0}^{\lfloor s/2\rfloor} \frac{1}{j!}\|L^j_{\chi_0^{(2)}} f_{2l+2j}^{(1,s-2j)}\|_{1-d_1-\delta_2}\\
&\le \sum_{j=0}^{\lfloor s/2\rfloor} \frac{1}{e}\left(\frac{\|X_0^{(2)}\|_{1-d_1}}{\delta_2^2\rho\sigma}+\frac{e|\zeta^{(2)}|}{\delta_2\rho}\right)^j \|f_{2l+2j}^{(1,s-2j)}\|_{1-d_1}\\
&\le \sum_{j=0}^{\lfloor s/2\rfloor} \frac{1}{e}\left(\frac{1}{\omega\delta_2^2\rho\sigma}+\frac{e}{4m\delta_2\rho^2}\right)^j \frac{1}{\omega^j}\nu_{1,2}^j\Xi_2^{3j}E^j\varepsilon^{2j}\nu_{1,s-2j}\Xi_2^{2(s-2j)}\frac{E}{2^{2l+2j}}\varepsilon^{s-2j}\\
&\le \nu_{2,s}^{(\mathrm{I})}\Xi_2^{2s}\frac{E}{2^{2l}}\varepsilon^s\ .
\end{aligned}$$

Coming to the second stage of the normalization step, the generating function $\chi_2^{(2)}$ is bounded by

$$\|\chi_2^{(2)}\|_{1-d_1-\delta_2} \le \frac{1}{\omega}\|f_2^{(\mathrm{I};1,2)}\|_{1-d_1-\delta_2} \le \frac{1}{\omega}3\nu_{1,2}\Xi_2^3\frac{E}{4}\varepsilon^2\ .$$

The terms $f_{2l}^{(2,s)}$ appearing in the expansion of the Hamiltonian $H^{(2)}$ are bounded as follows. For $s=1$ all the terms are unchanged, thus there is nothing to do. Furthermore, both $f_0^{(2,2)}$ and $f_2^{(2,2)}$ are trivially bounded with the norm of $f_0^{(\mathrm{I},1,2)}$ and $f_2^{(\mathrm{I},1,2)}$, respectively. Similarly to the first stage of the the normalization step, the terms $f_0^{(2,3)}$ and $f_0^{(2,4)}$ are bounded as follows

$$\|f_0^{(2,3)}\|_{1-d_2} \le \nu_{2,3}\Xi_2^5E\varepsilon^3\ ,\qquad \|f_0^{(2,4)}\|_{1-d_2} \le \nu_{2,4}\Xi_2^7E\varepsilon^4\ .$$

For the remaining terms one has

$$\|f_{2l}^{(2,s)}\|_{1-d_2} \le \sum_{j=0}^{[s/2]}\frac{1}{e}\left(\frac{1}{\delta_2^2\rho\sigma}\right)^j\frac{1}{\omega^j}(3\nu_{1,2})^j\Xi_2^{3j}\frac{E^j}{2^{2j}}\varepsilon^{2j}\nu_{2,s-2j}^{(\mathrm{I})}\Xi_2^{2s-4j}\frac{E}{2^{2l}}\varepsilon^{s-2j} \le \nu_{2,s}\Xi_2^{2s}\frac{E}{2^{2l}}\varepsilon^s\ .$$

This concludes the proof of the Lemma.

□

Lemma A.7 *Let $s=\lfloor s/r\rfloor r+m$, then for $0\le j\le\lfloor s/r\rfloor$ one has*

$$3rj-2j+b(r-1,s-jr,2l+2j)\le b(\mathrm{I};r-1,s,2l)\ .$$

Proof. The proof just requires a trivial computation, i.e.,

$$\begin{aligned}
&3rj - 2j + b(r-1, s-jr, 2l+2j) = \\
&\quad = 3rj - 2j + 3(s-jr) - \left\lfloor \frac{s-jr+r-2}{r-1} \right\rfloor - \left\lfloor \frac{s-jr+r-3}{r-1} \right\rfloor \\
&\quad = 3s - \left\lfloor \frac{s-j+r-2}{r-1} \right\rfloor - \left\lfloor \frac{s-j+r-3}{r-1} \right\rfloor \\
&\quad = 3s - \left\lfloor \frac{\lfloor s/r \rfloor r + m - j + r - 2}{r-1} \right\rfloor - \left\lfloor \frac{\lfloor s/r \rfloor r + m - j + r - 3}{r-1} \right\rfloor \\
&\quad = 3s - \left\lfloor \frac{s}{r} \right\rfloor - \left\lfloor \frac{\lfloor s/r \rfloor + m - j + r - 2}{r-1} \right\rfloor - \left\lfloor \frac{s}{r} \right\rfloor - \left\lfloor \frac{\lfloor s/r \rfloor + m - j + r - 3}{r-1} \right\rfloor \\
&\quad \leq 3s - \left\lfloor \frac{s+r-1}{r} \right\rfloor - \left\lfloor \frac{s+r-2}{r} \right\rfloor \\
&\quad \leq b(\mathrm{I}; r-1, s, 2l)
\end{aligned}$$

□

A.4.1 Proof of lemma 2.2

We proceed by induction. For $r = 1, 2$ just use Lemmas A.5 and A.6, respectively.

For $r > 2$, the estimates (26) for the generating functions follow directly from Lemma A.4, remarking that

$$b(r-1, r, 2) = b(\mathrm{I}; r-1, r, 2) = 3r - \left\lfloor \frac{2r-2}{r-1} \right\rfloor - \left\lfloor \frac{2r-3}{r-1} \right\rfloor = 3r - 3 \ .$$

The terms $f_{2l}^{(\mathrm{I};r-1,s)}$ appearing in the expansion of the Hamiltonian $H^{(\mathrm{I};r-1)}$ are bounded as follows. For $l = 0, 1$ and $s < r$ all the terms are unchanged, thus there is nothing to do. The term $f_0^{(\mathrm{I};r-1,r)}$ is trivially bounded with the norm of $f_0^{(r-1,r)}$. The term $f_2^{(\mathrm{I};r-1,r)}$ requires more care[5] since only the generating function $X_0^{(r)}$ plays a role and we get the following estimate

$$\|f_2^{(\mathrm{I};r-1,r)}\|_{1-d_{r-1}-\delta_r} \leq 3\nu_{r-1,r} \Xi_r^{3r-3} \frac{E}{4} \varepsilon^r \ .$$

For $l = 0$ and $r < s \leq 2r$,

$$\begin{aligned}
\|f_0^{(\mathrm{I};r-1,s)}\|_{1-d_{r-1}-\delta_r} &\leq \|f_0^{(r-1,s)}\|_{1-d_{r-1}-\delta_r} + \|L_{\chi_0^{(r)}} f_2^{(r-1,s-r)}\|_{1-d_{r-1}-\delta_r} \\
&\leq \nu_{r-1,s} \Xi_r^{b(r-1,s,0)} E \varepsilon^s \\
&\qquad + \frac{1}{e} \left(\frac{\|X_0^{(r)}\|_{1-d_{r-1}}}{\delta_r^2 \rho \sigma} + \frac{e|\zeta^{(r)}|}{\delta_r \rho} \right) \nu_{r-1,s-r} \Xi_r^{b(r-1,s-r,2)} \frac{E}{2^2} \varepsilon^{s-r} \\
&\leq \nu_{r-1,s} \Xi_r^{b(r-1,s,0)} E \varepsilon^s \\
&\qquad + \frac{1}{e} \left(\frac{E}{\omega \delta_r^2 \rho \sigma} + \frac{eE}{4m\delta_r \rho} \right) \nu_{r-1,r} \Xi_r^{b(r-1,r,2)} \varepsilon^r \nu_{r-1,s-r} \Xi_r^{b(r-1,s-r,2)} \frac{E}{2^2} \varepsilon^{s-r} \\
&\leq \Xi_r^{b(\mathrm{I};r-1,s,0)} \nu_{r,s}^{(\mathrm{I})} E \varepsilon^s \ ,
\end{aligned}$$

[5]See the proofs of Lemma A.6 and Lemma A.4.

where we have used the trivial inequality

$$3r-2+b(r-1,s-r,2)\le b(\mathrm{I};r-1,s,0)\ .$$

For the remaining terms one has

$$\begin{aligned}\|f_{2l}^{(\mathrm{I};r-1,s)}\|_{1-d_{r-1}-\delta_r} &\le \sum_{j=0}^{\lfloor s/r\rfloor}\frac{1}{j!}\|L^j_{\chi_0^{(r)}}f_{2l+2j}^{(r-1,s-jr)}\|_{1-d_{r-1}-\delta_r}\\
&\le \sum_{j=0}^{\lfloor s/r\rfloor}\frac{1}{e}\left(\frac{\|X_0^{(r)}\|_{1-d_{r-1}}}{\delta_r^2\rho\sigma}+\frac{e|\zeta^{(r)}|}{\delta_r\rho}\right)^j\nu_{r-1,s-jr}\Xi_r^{b(r-1,s-jr,2l+2j)}\frac{E}{2^{2l+2j}}\varepsilon^{s-jr}\\
&\le \sum_{j=0}^{\lfloor s/r\rfloor}\frac{1}{e}\left(\frac{E}{\omega\delta_r^2\rho\sigma}+\frac{eE}{4m\delta_r\rho}\right)^j\nu_{r-1,r}^j\Xi_r^{b(r-1,r,2)j}\varepsilon^{rj}\\
&\qquad\qquad\times\nu_{r-1,s-jr}\Xi_r^{b(r-1,s-jr,2l+2j)}\frac{E}{2^{2l+2j}}\varepsilon^{s-jr}\\
&\le \nu_{r,s}^{(\mathrm{I})}\Xi_r^{b(\mathrm{I};r-1,s,2l)}\frac{E}{2^{2l}}\varepsilon^s\ ,\end{aligned}$$

where we have used the trivial inequality

$$3rj-2j+b(r-1,s-jr,2l+2j)\le b(\mathrm{I};r-1,s,2l)\ .$$

Coming to the second stage of the normalization step, just notice that the bound for the generating function $\chi_2^{(r)}$ is similar to the one of $\chi_0^{(r)}$ and in particular it has exactly the same exponent for the coefficient Ξ_r. Thus all the estimates appearing in the expansion of the Hamiltonian $H^{(r)}$ are nothing but a minor *variazione*, mutatis mutandis, with respect to the first stage of the normalization step. This concludes the proof of the Lemma.

B Proof of Proposition 3.1

The Proposition is a direct consequence of the Contraction Principle applied to a suitable closed ball centered in x_0. Indeed, by following a standard procedure (see, i.e., [20]), let us formulate the original problem as a fixed point problem, namely $\Upsilon(x,\varepsilon)=0$ if and only if $A(x,\varepsilon)=x$, where

$$A(x,\varepsilon)=x-[\Upsilon'(x_0,\varepsilon)]^{-1}\Upsilon(x,\varepsilon)\ .$$

We first of all show that A is a contraction of a sufficiently small ball centered in x_0. We first rewrite our assumptions in a more general form

$$\|\Upsilon(x_0,\varepsilon)\|\le\mu\ ,\qquad \|[\Upsilon'(x_0,\varepsilon)]^{-1}\|_{\mathcal{L}(V)}\le M\ ,$$

and we introduce the auxiliary quantities

$$\eta=M\mu=C_1C_2|\varepsilon|^{\beta-\alpha}\ ,\qquad\qquad h=MC_3\eta=C_1C_2^2C_3|\varepsilon|^{\beta-2\alpha}\ .$$

Notice that the condition $\beta>2\alpha$ is necessary in order to have

$$\lim_{\varepsilon\to0}h=0\ .$$

The main ingredient is the continuity of Υ', since $\Upsilon \in \mathcal{C}^1$ locally around x_0 (independently from ε). From finite increment formula we get, for $x, y \in B(x_0, r) \subset \mathcal{U}(x_0)$

$$\|A(x,\varepsilon) - A(y,\varepsilon)\| \leq \left(\sup_{z \in B(x_0,r)} \left\|A'(z,\varepsilon)\right\|_{\mathcal{L}(V)} \right) \|x - y\| \ ;$$

thus, we aim at showing that, with a suitable choice of the radius r, we have

$$\sup_{z \in B(x_0,r)} \left\|A'(z,\varepsilon)\right\|_{\mathcal{L}(V)} < 1 \ .$$

Since

$$A'(z,\varepsilon) = \mathbb{I} - \left[\Upsilon'(x_0,\varepsilon)\right]^{-1} \Upsilon'(z,\varepsilon) = \left[\Upsilon'(x_0,\varepsilon)\right]^{-1} \left[\Upsilon'(x_0,\varepsilon) - \Upsilon'(z,\varepsilon)\right]$$

we get

$$\begin{aligned} \left\|A'(z,\varepsilon)\right\|_{\mathcal{L}(V)} &\leq \left\| \left[\Upsilon'(x_0,\varepsilon)\right]^{-1} \right\|_{\mathcal{L}(V)} \left\|\Upsilon'(x_0,\varepsilon) - \Upsilon'(z,\varepsilon)\right\|_{\mathcal{L}(V)} \leq \\ &\leq M \left\|\Upsilon'(x_0,\varepsilon) - \Upsilon'(z,\varepsilon)\right\|_{\mathcal{L}(V)} \ . \end{aligned}$$

From the continuity of Υ' it follows that, provided $\|z - x_0\|$ is small enough, it is possible to make $\Upsilon'(x_0,\varepsilon) - \Upsilon'(z,\varepsilon)$ arbitrary small. The Lipschitz-continuity estimate[6] in the hypotheses of the Proposition allows to explicitly deal with this issue. Indeed, from

$$\left\|\Upsilon'(x_0,\varepsilon) - \Upsilon'(z,\varepsilon)\right\|_{\mathcal{L}(V)} \leq C_3 \|z - x_0\| \ ,$$

we get

$$\left\|A'(z,\varepsilon)\right\|_{\mathcal{L}(V)} \leq MC_3 \|z - x_0\| \leq MC_3 r =: q \ , \qquad \forall z \in B(x_0, r) \ ,$$

and also

$$\sup_{z \in B(x_0,r)} \left\|A'(z,\varepsilon)\right\|_{\mathcal{L}(V)} \leq q \ .$$

In order to show that $\Upsilon(B(x_0,r)) \subset B(x_0,r)$, namely that $\|z - x_0\| \leq r$ implies $\|A(z,\varepsilon) - x_0\| \leq r$, we start splitting

$$\|A(z,\varepsilon) - x_0\| \leq \|A(z,\varepsilon) - A(x_0,\varepsilon)\| + \|A(x_0,\varepsilon) - x_0\| \ .$$

We will separately estimate the two r.h.t.. From the bound on $A'(z,\varepsilon)$ we get

$$\|A(z,\varepsilon) - A(x_0,\varepsilon)\| \leq \sup_{z \in B(x_0,r)} \left\|A'(z,\varepsilon)\right\|_{\mathcal{L}(V)} \|z - x_0\| \leq qr \ .$$

on the other hand, by exploiting the initial definition of $A(x,\varepsilon)$, one has

$$\begin{aligned} \|A(x_0,\varepsilon) - x_0\| &= \left\|x_0 - [\Upsilon'(x_0,\varepsilon)]^{-1}\Upsilon(x_0,\varepsilon) - x_0\right\| = \left\|[\Upsilon'(x_0,\varepsilon)]^{-1}\Upsilon(x_0,\varepsilon)\right\| \leq \\ &\leq \left\| \left[\Upsilon'(x_0,\varepsilon)\right]^{-1} \right\|_{\mathcal{L}(V)} \|\Upsilon(x_0,\varepsilon)\| \leq M\mu \ . \end{aligned}$$

Hence, in order to have $\Upsilon(B(x_0,r)) \subset B(x_0,r)$, it must happen

$$M\mu + qr \leq r \ .$$

[6] Actually Holder-continuity will be sufficient, modifying the conditions on α and β.

Thus, two independent conditions have to be satisfied:

$$MC_3 r < 1 \ , \qquad \eta + MC_3 r^2 \le r \ .$$

The second is equivalent to

$$MC_3 r^2 - r + \eta \le 0 \ ,$$

which can be re-scaled to

$$r = \eta\rho \ , \qquad\qquad h\rho^2 - \rho + 1 \le 0 \ .$$

The corresponding equation, under the condition $h < \frac{1}{4}$, has the two zeros

$$t_\pm = \frac{1}{2h}\Big(1 \pm \sqrt{1-4h}\Big) \ .$$

Moreover one has $t_- < 2$, since $1-4h < \sqrt{1-4h}$, and for $h \sim 0$ we get $t_-(h) \sim 1$. Collecting the above information, the radius r has to fulfill

$$\eta t_- \le r \le t_+\eta \ .$$

If we make the more restrictive choice

$$\eta t_- \le r \le 2\eta \ ,$$

then, from $h < \frac{1}{4}$, it follows that Υ is an $\frac{1}{2}$-contraction map

$$MC_3 r < 2MC_3\eta = 2h < \frac{1}{2} \ .$$

In our case, $h < \frac{1}{4}$ comes directly from being $h(\varepsilon)$ infinitesimal w.r.t. ε; thus for ε small enough the condition is satisfied. Moreover, from $h(\varepsilon) \approx 1$, one deduces that the optimal choice for the radius is

$$r(\varepsilon) = \eta t_- \approx C_1 C_2 |\varepsilon|^{\beta-\alpha} \ .$$

□

Remark B.1 *The above Proposition shows that x_0 is a better approximation of the true solution as α decreases, which means as the differential $\Upsilon'(x_0,\varepsilon)$ is bounded independently on ε*

$$\left\|\Upsilon'(x_0,\varepsilon)\right\| \ge C \qquad \Rightarrow \qquad \alpha = 0 \ .$$

At the limiting case $\alpha = 0$, it is possible to choose $r = \mathcal{O}(\varepsilon^\beta)$.

Acknowledgments

We warmly thank Simone Paleari for useful discussions on the normal form scheme and the two anonymous referees for helpful suggestions concerning the presentation of this paper and for pointing out the connections with some recent results on highly degenerate Hamiltonian systems. T.P. also thanks his former Master student Letizia Zanni, for her introductory work on Moser-Weinstein approach. T.P. and M.S. have been partially supported by the National Group of Mathematical Physics (GNFM-INdAM), as part of the project *"Normal Form techniques in Lattice Dynamics and Celestial Mechanics: perturbed resonant dynamics via resonant normal forms"*.

References

[1] T. Ahn. Multisite oscillations in networks of weakly coupled autonomous oscillators. *Nonlinearity*, 11(4):965–989, 1998.

[2] S. Aubry. Breathers in nonlinear lattices: existence, linear stability and quantization. *Phys. D*, 103(1-4):201–250, 1997.

[3] G. Benettin, L. Galgani, A. Giorgilli, J.-M. Strelcyn. A proof of Kolmogorov's theorem on invariant tori using canonical transformations defined by the Lie method. *Nuovo Cimento B (11)*, 79(2):201–223, 1984.

[4] M. Berti. *Nonlinear oscillations of Hamiltonian PDEs.* Progress in Nonlinear Differential Equations and their Applications, 74. Birkhäuser Boston, Inc., Boston, MA, 2007.

[5] L. Biasco, L. Chierchia, E. Valdinoci. N-dimensional elliptic invariant tori for the planar $(N+1)$-body problem. *SIAM J. Math. Anal.*, 37(5):1560–1588, 2006.

[6] C.-Q. Cheng. Birkhoff-Kolmogorov-Arnold-Moser tori in convex Hamiltonian systems. *Comm. Math. Phys.*, 177:529–559, 1996.

[7] L. Corsi, R. Feola, G. Gentile. Lower-dimensional invariant tori for perturbations of a class of non-convex Hamiltonian functions. *J. Stat. Phys.*, 150(1):156–180, 2013.

[8] L. Corsi, G. Gentile. Resonant motions in the presence of degeneracies for quasi-periodically perturbed systems. *Ergodic Theory Dynam. Systems*, 35(4):1079–1140, 2015.

[9] L.H. Eliasson Perturbations of stable invariant tori for Hamiltonian systems. *Ann. Scuola Norm. Sup. Pisa, Cl. Sci., IV Serie*, 15:115–147, 1988.

[10] A. Giorgilli, U. Locatelli. Kolmogorov theorem and classical perturbation theory. *Zeitschrift für angewandte Mathematik und Physik ZAMP*, 48(2):220–261, Mar 1997.

[11] A. Giorgilli, M. Sansottera. Methods of algebraic manipulation in perturbation theory. *Workshop Series of the Asociacion Argentina de Astronomia*, 3:147–183, 2011.

[12] A. Giorgilli. Exponential stability of Hamiltonian systems. *Dynamical systems. Part I*, Pubbl. Cent. Ric. Mat. Ennio Giorgi, pages 87–198. Scuola Norm. Sup., Pisa, 2003.

[13] A. Giorgilli. On a theorem of Lyapounov. *Rend. Ist. Lomb. Acc. Sc. Lett.*, 146:133–160, 2012.

[14] A. Giorgilli, U. Locatelli, M. Sansottera. On the convergence of an algorithm constructing the normal form for elliptic lower dimensional tori in planetary systems. *CeMDA*, 119(3-4):397–424, 2014.

[15] A. Giorgilli, U. Locatelli, M. Sansottera. Improved convergence estimates for the Schroder-Siegel problem *Ann. Mat. Pura ed Applicata*, 194:995–1013, (2015).

[16] W. Gröbner. *Die Lie-Reihen und ihre Anwendungen*, volume 3 of *Mathematische Monographien.* VEB Deutscher Verlag der Wissenschaften, Berlin, 1960.

[17] Y. Han, Y. Li, Y. Yi. Invariant tori in Hamiltonian systems with high order proper degeneracy. *Ann. Henri Poincaré*, 10:1419–1436, 2010.

[18] D. Hennig, G.P. Tsironis. Wave transmission in nonlinear lattices. *Phys. Rep.*, 307(5-6):333–432, 1999.

[19] P.G. Kevrekidis. *The discrete nonlinear Schrödinger equation*, volume 232 of *Springer Tracts in Modern Physics*. Springer-Verlag, Berlin, 2009.

[20] A.N. Kolmogorov, S.V. Fomin. *Elementy teorii funktsii i funktsionalnogo analiza.* "Nauka", Moscow, sixth edition, 1989. With a supplement, "Banach algebras", by V. M. Tikhomirov.

[21] V. Koukouloyannis. Non-existence of phase-shift breathers in one-dimensional Klein-Gordon lattices with nearest-neighbor interactions. *Phys. Lett. A*, 377(34-36):2022–2026, 2013.

[22] V. Koukouloyannis, P.G. Kevrekidis. On the stability of multibreathers in Klein-Gordon chains. *Nonlinearity*, 22(9):2269–2285, 2009.

[23] S.B. Kuksin. The perturbation theory for the quasi-periodic solutions of infinitedimensional Hamiltonian systems and its applications to the Korteweg de Vries equation. *Matem. Sbornik*, 136, 1988. English transl. in Math. USSR Sbornik, 64:397–413, 1989.

[24] Y. Li, Y. Yi. A quasi-periodic Poincaré's theorem *Mathematische Annalen*, 326:649–690, 2003.

[25] E. Meletlidou, G. Stagika. On the continuation of degenerate periodic orbits in Hamiltonian systems. *RCD*, 11(1):131–138, 2006.

[26] V.K. Melnikov. On some cases of preservation of quasiperiodic motions under a small perturbation of the Hamiltonian. *Dokl. Akad. Nauk SSSR*, 165:1245–1248, 1965. English transl. in Soviet Math. Dokl., 6, 1965.

[27] V.K. Melnikov. On a family of quasiperiodic solutions of a Hamiltonian system *Dokl. Akad. Nauk SSSR*, 181:546–549, 1968. English transl. in Soviet Math. Dokl. 9, 1968.

[28] J. Moser. Periodic orbits near an equilibrium and a theorem by Alan Weinstein. *Comm. Pure Appl. Math.*, 29(6):724–747, 1976.

[29] J. Moser. Convergent series expansions for quasi-periodic motions. *Math. Ann*, 169:136–176, 1967.

[30] D.E. Pelinovsky, P.G. Kevrekidis, D.J. Frantzeskakis. Persistence and stability of discrete vortices in nonlinear Schrödinger lattices. *Phys. D*, 212(1-2):20–53, 2005.

[31] D. Pelinovsky, A. Sakovich. Multi-site breathers in Klein-Gordon lattices: stability, resonances and bifurcations. *Nonlinearity*, 25(12):3423–3451, 2012.

[32] T. Penati, V. Koukouloyannis, M. Sansottera, P.G. Kevrekidis, S. Paleari. On the nonexistence of degenerate phase-shift multibreathers in a zigzag Klein-Gordon model via dNLS approximation. *preprint*, 2017.

[33] T. Penati, M. Sansottera, S. Paleari, V. Koukouloyannis, P.G. Kevrekidis. On the nonexistence of degenerate phase-shift discrete solitons in a dNLS nonlocal lattice. *Physica D*, 2017. doi:10.1016/j.physd.2017.12.012

[34] J. Pöschel. On elliptic lower dimensional tori in Hamiltonian systems. *Math. Z.*, 202:559–608, 1989.

[35] H. Poincaré. *Les méthodes nouvelles de la mécanique céleste. Tome I. Solutions périodiques. Non-existence des intégrales uniformes. Solutions asymptotiques.* Dover Publications, Inc., New York, N.Y., 1957.

[36] H. Poincaré. *Œuvres. Tome VII.* Les Grands Classiques Gauthier-Villars. [Gauthier-Villars Great Classics]. Éditions Jacques Gabay, Sceaux, 1996.

[37] M. Sansottera, M. Ceccaroni. Rigorous estimates for the relegation algorithm. *CeMDA*, 127(1):1–18, Jan 2017.

[38] M. Sansottera, U. Locatelli, A. Giorgilli. A semi-analytic algorithm for constructing lower dimensional elliptic tori in planetary systems. *CeMDA*, 111(3):337–361, 2011.

[39] D.V. Treshchev. The mechanism of destruciton of resonant tori of Hamiltonian systems. *Math. USSR Sb.*, 68:181–203, 1991.

[40] G. Voyatzis, S. Ichtiaroglou. Degenerate bifurcations of resonant tori in Hamiltonian systems. *Internat. J. Bifur. Chaos Appl. Sci. Engrg.*, 9(5):849–863, 1999.

[41] A. Weinstein. Normal modes for nonlinear Hamiltonian systems. *Invent. Math.*, 20:47–57, 1973.

[42] L. Xu, Y. Li, Y. Yi. Lower-Dimensional Tori in Multi-Scale, Nearly Integrable Hamiltonian Systems *Ann. Henri Poincaré*, 18:53–83, 2017.

[43] L. Xu, Y. Li, Y. Yi. Poincaré-Treshchev Mechanism in Multi-scale, Nearly Integrable Hamiltonian Systems *J. Nonlin. Sc.*, 28(1):337–369, 2018.

[44] V.A. Yakubovich, V.M. Starzhinskii. *Linear differential equations with periodic coefficients. 1, 2.* Halsted Press [John Wiley & Sons] New York-Toronto, Ont.; Israel Program for Scientific Translations, Jerusalem-London, 1975. Translated from Russian by D. Louvish.

[45] E. Zehnder. An implicit function theorem for small divisor problems, *Bull. Amer. Math. Soc.*, 80:174–179, 1974.

[46] E. Zehnder. Generalized implicit function theorems with applications to some small divisor problems. I, *Comm. Pure Appl. Math.*, 28:91–140, 1975.

[47] E. Zehnder. Generalized implicit function theorems with applications to some small divisor problems. II, *Comm. Pure Appl. Math.*, 29:49–111, 1976.